



HAL
open science

Caractérisation de l'énergie noire : développement d'une nouvelle génération de modélisation spectrophotométrique empirique des supernovæ de type Ia

Guy Augarde

► To cite this version:

Guy Augarde. Caractérisation de l'énergie noire : développement d'une nouvelle génération de modélisation spectrophotométrique empirique des supernovæ de type Ia. Astrophysique [astro-ph]. Sorbonne Université, 2022. Français. NNT : 2022SORUS343 . tel-03927516

HAL Id: tel-03927516

<https://theses.hal.science/tel-03927516v1>

Submitted on 6 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Sorbonne Université

École doctorale n° 560 : Sciences de la Terre et de l'Environnement et Physique de l'Univers de Paris (STEP'UP)

Laboratoire de Physique Nucléaire et de Hautes Energie (LPNHE)

Caractérisation de l'énergie noire : développement d'une nouvelle génération de modélisation spectrophotométrique empirique des supernovæ de type Ia.

Par : Guy AUGARDE

Thèse de doctorat de Cosmologie

Présenté et soutenue publiquement le 20 septembre 2022

Dirigée par : Nicolas REGNAULT

Devant un jury composé de :

M. Julien Guy,	Rapporteur
M. Richard Kessler,	Rapporteur
Mme. Vanina Ruhlmann-Kleider,	Examinatrice
Mme. Delphine Hardin,	Examinatrice
M. Mickael Rigault,	Examinateur
M. Marc Betoule,	Invité
M. Nicolas Regnault,	Directeur de thèse

Remerciements

Ce manuscrit est la somme de trois années de travail intense, durant lesquelles beaucoup de personnes m'ont aidé et soutenu. Je souhaite ici les remercier sincèrement.

J'aimerais tout particulièrement remercier le jury de cette thèse : Julien Guy, d'abord pour ton développement des modèles SALT, pour avoir rapporté mon travail et enfin pour tes commentaires qui ont grandement amélioré la qualité du discours. Richard Kessler ensuite pour avoir pris le temps de rapporter mon manuscrit *en français* et pour avoir contribué aussi par tes commentaires à l'éclaircissement des messages. Vanina Ruhlmann-Kleider, je tiens à te remercier pour ta lecture détaillée du manuscrit et pour tes commentaires. Merci Delphine Hardin pour ta lecture du manuscrit, tous tes commentaires et surtout ton suivi depuis le Master NPAC. Enfin merci Mickael Rigault pour ta lecture du manuscrit et tes questions pertinentes.

Je souhaite remercier l'école doctorale STEP'UP et l'ancienne Université Pierre et Marie Curie, nouvelle Sorbonne Université et tous les professeurs que j'y ai rencontrés durant ces huit années d'études, pour m'avoir permis de réaliser ma thèse. Merci notamment à Micheal Joyce qui aura été mon premier professeur de physique à l'Université en L1 et mon dernier en M2. Je souhaite aussi remercier les membres de mon comité de suivi de thèse. Merci Philippe Gris pour ton suivi du déroulement de cette thèse avec tant d'attention. Je tiens aussi à remercier Olivier Martineau, pour ton parrainage, nos discussions, tes encouragements et pour ta sincérité.

Je remercie le LPNHE pour m'avoir accueilli pendant ces trois années et particulièrement le service informatique qui a tout fait pour nous permettre de travailler à distance durant les différents confinements sans rencontrer aucun problème. En réalité, mon histoire avec le LPNHE est plus ancienne, ma classe de Terminale avait eu la chance en 2014, de visiter le laboratoire dans le cadre de la fête de la Science. Je souhaite donc remercier les organisateurs et les participants de cette fête, car ma thèse en est le fruit. Ce passage avait éveillé ma curiosité et mon intérêt pour la physique et on voit jusqu'où cela m'a mené !

Merci au groupe de cosmologie pour m'avoir si bien intégré, merci Pierre Antilogous pour ta bonne humeur, merci Pierre Astier pour ta vision du monde, merci Jérémy pour m'avoir aidé à améliorer la soutenance et pour tes encouragements, merci Pierre-Francois pour ton aide et tes conseils, merci Sylvain d'avoir toujours été disponible pour répondre à mes questions, merci Laurent pour toutes tes histoires passionnantes... Merci à tous pour ces discussions et tous vos enseignements.

Je tiens aussi à remercier chaleureusement le sous groupe SNe-NaCl : d'abord, Sébastien pour ton enseignement, ta disponibilité, ta bonne humeur et pour toutes nos discussions matinales en passant devant mon bureau. Merci Marc pour ta pédagogie, ta considération et ta curiosité. Je continuerai à faire plein de petits jouets inutiles en pensant à toi. Merci à vous deux pour vos encouragements, votre soutien et votre aide, merci d'avoir pris le temps de discuter de Physique et surtout d'autres choses même en plein coeur de la tempête. Merci pour toutes les réunions du lundi matin qui vont, je vous l'assure, beaucoup me manquer ; pour vos contributions à ce travail, pour tous vos conseils d'expression orale, qui comme vous le savez est l'exercice que je préfère et bien sûr merci pour les relectures du manuscrit. Je souhaite aussi remercier Leander et Dylan et bon courage pour la suite !

Enfin je te remercie tout particulièrement Nicolas. Merci pour la qualité et la générosité de ton enseignement d'abord en tant que professeur de cosmologie durant l'année de NPAC, ensuite pendant le stage de pré-thèse mais surtout durant ces trois années de thèse. Merci d'avoir tant de fois relu et corrigé mon code (pas toujours très clair. . .) et d'avoir pris le temps de m'expliquer les règles pour le pythonifier ! Merci d'avoir été aussi disponible, d'avoir pris le temps de discuter tard le soir ou tôt le week-end et durant les temps assez angoissants des différents confinements. Merci de m'avoir accompagné lorsque la bombe SALT3 a été publiée au milieu de la thèse. Merci de m'avoir fait confiance pour encadrer le stage de Timothé Alezraa. J'en profite pour te remercier Timothé pour ton travail d'exploration du lot d'entraînement de Salt3, ce n'est pas le travail le plus facile et tu l'as parfaitement réalisé. Merci Nicolas d'avoir pris le temps de discuter de questions plus personnelles et d'avoir toujours eu une vision honnête des réalités. Merci d'avoir attentivement relu ce manuscrit et même en plein coeur de l'été. Merci pour cela et pour le reste, si je devais tout lister ce serait sûrement aussi long que le contenu de cette thèse, qui est par elle-même le meilleur des remerciements. Je souhaite sincèrement à tous les doctorants d'avoir un tel encadrement.

Je souhaite remercier tendrement ma famille, mes parents et mon frère à qui je dois tant, pour leur amour et leur soutien. Merci pour les valeurs que vous m'avez transmises.

Merci à tous mes amis, cette famille que l'on choisit, merci la sdk pour ces quinze années de rigolades, de soutien et surtout de conneries. Je vous remercie de m'avoir vidé l'esprit et sorti la tête du guidon. Merci à mes co-bureau Viet, Sara et Lucile pour toutes ces discussions. Je remercie aussi tous les autres thésards du LPNHE pour tous les bons moments passés ensemble.

Je tiens à remercier une des personnes les plus impliquées à l'aboutissement de ce travail et dont le nom n'apparaît nulle part. Merci Océane pour ton soutien quotidien et indéfectible. Merci pour ton enthousiasme pour mon travail, pour tes encouragements essentiels, pour m'avoir fait répéter *toutes* mes présentations et pour avoir relu et corrigé les fautes d'orthographe de ce manuscrit. Notre couple a beaucoup grandi pendant cette thèse ; le premier étant la stabilité indispensable à la réussite du second.

Merci à la France, cette vieille dame qui permet à un petit gars de réaliser 26 ans d'études dans des conditions presque idéales. Je veux aussi remercier de tout coeur Charles Aznavour, Serge Reggiani, Barbara et tant d'autres pour ces superbes musiques et sans qui, la route aurait été un peu plus longue.

Enfin je vous remercie par avance, lecteurs et lectrices de ce manuscrit, j'espère que vous y trouverez ce que vous cherchez !

Table des matières

Introduction	9
I Caractérisation de l'énergie noire avec les Supernovæ de type Ia	11
1 Le modèle standard de la Cosmologie	13
1.1 Les équations d'Einstein	15
1.2 Un univers isotrope, homogène et en expansion	15
1.2.1 Isotropie	15
1.2.2 Homogénéité	15
1.2.3 Loi de Hubble-Lemaître	15
1.2.4 Métrique FRLW	17
1.3 Observables	18
1.3.1 Décalage spectral ou <i>Redshift</i>	18
1.3.2 Les distances cosmologiques	19
1.4 Les équations de Friedmann	22
1.5 Le modèle Λ CDM	25
1.5.1 Le contenu énergétique de l'Univers	25
1.5.2 Le fond diffus cosmologique	32
1.5.3 Croissance des structures	33
1.6 L'étude de l'énergie noire avec la combinaison de différentes sondes cosmologiques	34
1.6.1 Les chandelles standards : SNe Ia	34
1.6.2 Le pic de l'oscillation acoustique des baryons	35
1.6.3 Combinaison de sondes	39
1.7 Autres utilisations cosmologiques des SNe Ia	41
1.7.1 Diagramme de Hubble pour mesurer H_0	41
1.7.2 Vitesses particulières	44
1.8 Succès & tensions du modèle	45
2 Les Supernovæ de type Ia	49
2.1 Observations et mesures des supernovæ	51
2.2 Les Supernovæ (SNe)	51
2.2.1 Classification	52
2.2.2 Proportions & taux d'explosion	54
2.2.3 Explosion par effondrement gravitationnel	55

2.2.4	Supernova thermonucléaire	56
2.3	Propriétés observationnelles des SNe Ia	58
2.3.1	Propriétés spectrales	58
2.3.2	Propriétés photométriques	60
2.3.3	Standardisation	61
2.3.4	Propriétés de la galaxie hôte	64
2.3.5	Sous-classes	66
2.4	Conclusion	67
3	Observation de SNe pour une analyse cosmologique	69
3.1	Les instruments de mesure	71
3.1.1	Une première révolution : le capteur CCD	71
3.1.2	Une seconde révolution : les caméras grand champ	73
3.1.3	Observations spectroscopiques	75
3.2	Données de suivi	77
3.2.1	Photométrie en bande large	78
3.2.2	Analyse des données photométriques	79
3.2.3	Données spectroscopiques	83
3.3	Les relevés	84
3.3.1	Première génération	84
3.3.2	Génération 1.5	86
3.3.3	Seconde génération	87
3.3.4	Troisième génération	88
3.3.5	Futurs relevés	89
3.4	Diagramme de Hubble - contraintes sur l'équation d'état de l'énergie noire	90
3.4.1	Mesure de distance avec des SNe Ia	90
3.4.2	Modèle empirique d'ajustement de courbes de lumière	91
3.4.3	Diagramme de Hubble	91
3.5	Sources d'incertitudes	92
3.5.1	Incetitudes statistiques	93
3.5.2	Incetitudes systématiques	94
3.6	Conclusion	96
 II Développement d'une nouvelle génération de modélisation spectrophotométrique empirique des Supernovæ de type Ia :		
<i>Nouveaux algorithmes de Courbes de lumière (NaCl)</i>		99
4	Modèles empiriques de flux de Supernovæ de type Ia	101
4.1	Introduction	102
4.2	Modèle d'ajustement de courbes de lumière	103
4.2.1	L'approche historique	103
4.2.2	MLCS/MLCS2k2	105
4.3	Modèle de distribution en énergie spectrale	105
4.3.1	SALT	106
4.3.2	SiFTO	106
4.3.3	SALT2	107
4.4	Les modèles de la SNfactory	109
4.4.1	SNEMO	109
4.4.2	SUGAR	111

4.4.3	Twins Embedding	111
4.5	Deux nouveaux ré-entraînements de SALT2	113
4.6	Conclusion	115
5	Le modèle spectrophotométrique	117
5.1	De la luminosité de la SN aux données de suivi	119
5.2	Le modèle SALT2	120
5.2.1	Modèle spectroscopique	123
5.2.2	La modélisation photométrique.	123
5.2.3	Les contraintes	124
5.2.4	Recalibration des spectres	125
5.2.5	La régularisation	127
5.3	Modèle d'erreur	127
5.3.1	<i>Error Snake</i>	128
5.3.2	Modélisation de la dispersion en couleur : <i>Color Scatter</i>	128
5.4	Propagation des incertitudes de calibration	129
5.5	Des équations au code	130
5.5.1	Implémentation des splines	131
5.5.2	Nombre de paramètres	132
5.5.3	Évaluation des flux photométriques	133
5.6	Entraînement du modèle	135
5.7	Conclusion	136
6	Minimiseur	143
6.1	Estimateur par maximum de vraisemblance.	144
6.2	Algorithme de score de Fisher : ajustement d'un modèle non linéaire	145
6.2.1	... sous contraintes	147
6.2.2	... avec de la régularisation	147
6.2.3	... en propageant les incertitudes de calibration	150
6.2.4	... et estimation de modèles d'erreur simultanément	151
6.2.5	χ^2 total	152
6.3	Les équations normales	153
6.3.1	Dérivées de l'ajustement du modèle	153
6.3.2	Dérivé du χ_{cons}^2	154
6.3.3	Dérivé du χ_{Reg}^2	155
6.3.4	Dérivé du χ_{calib}^2	155
6.3.5	Dérivé du χ_{CS}^2	155
6.3.6	L'équation finale	156
6.4	Estimation des incertitudes sur les paramètres	157
6.5	Conclusion	158
7	Simulations	159
7.1	Simulation de différents lot d'entraînement	160
7.1.1	Simulations idéales	160
7.1.2	Lots de type LSST	162
7.1.3	Simulation lots JLA et K21	163
7.2	Flux	171
7.3	Génération du bruit ajouté	172
7.3.1	Exemple SNe de JLA	173
7.4	Conclusion	174

8	Procédure de minimisation	175
8.1	Procédure de minimisation sur les simulations LSST	176
8.1.1	État initial	176
8.1.2	Critères de convergence	177
8.1.3	Détermination des hyper-paramètres	178
8.2	Procédure d'entraînement	179
8.2.1	Premier ajustement d'initialisation	179
8.2.2	Ajustement du modèle d'erreur	183
8.2.3	Fit complet	183
8.3	Temps de calcul	188
8.4	Test sur lot de données réalistes	190
8.4.1	JLA	190
8.4.2	K21	199
8.5	Conclusion	199
9	Études des biais et de l'efficacité de l'estimateur	201
9.1	Description de la méthode	202
9.1.1	Modules de distance	202
9.1.2	Incertitudes sur les modules de distance	203
9.2	Biais	203
9.3	Variance	210
9.3.1	Tests sur lots de données LSST-like	210
9.3.2	JLA	218
9.4	Conclusion	218
	Conclusion	225
A	Dérivées	227
A.1	Modèle	227
A.2	Modèles d'erreur	228
	Table des figures	229
	Liste des tableaux	239
	Bibliographie	240

Introduction

La découverte de l'accélération de l'expansion de l'Univers au tournant du millénaire est, avec les premières mesures de précision de la constante de Hubble et la mesure de la taille angulaire du pic acoustique du CMB, l'une des trois observations cruciales qui ont permis l'avènement du modèle standard de la cosmologie (Λ CDM). Avec seulement 6 paramètres, Λ CDM fournit un cadre explicatif cohérent à toutes les observables de l'Univers proche et lointain. En revanche, les causes exactes de l'accélération cosmique sont inconnues. Deux hypothèses sont couramment proposées : l'énergie noire, un fluide répulsif de nature inconnue et dominant la densité totale de l'Univers aux grandes échelles, ou simplement le fait que la relativité générale est incorrecte aux échelles cosmologiques. Des indices sur la nature de l'énergie noire peuvent être obtenus via une mesure fine de l'histoire de l'expansion de l'Univers, qui est une fonction de son équation d'état : $w = p/\rho$. Par ailleurs, une fois l'histoire de l'expansion connue, celle de la formation des structures est une prédiction directe de la relativité générale. Ces observations sont au programme de la décennie en cours avec trois grands relevés sols (ZTF, LSST et DESI) et deux relevés spatiaux (*Euclid* et *Roman*).

Les supernovæ de type Ia (SNe Ia) sont d'excellents indicateurs de distance de luminosité aux échelles cosmologiques. Elles présentent une uniformité de leur luminosité au pic de l'ordre de 40%, réductible à 14% en exploitant plusieurs corrélations empiriques entre la luminosité et d'autres observables photométriques : la couleur et la durée de l'événement notamment. Historiquement, c'est la mesure de la relation entre distance de luminosité des SNe et redshift qui a permis la découverte de l'accélération de l'expansion de l'Univers par les équipes du *High-Z Supernova Search Team* et du *Supernova Cosmology Project*. Un effort soutenu au cours des deux décennies qui ont suivi, a permis d'améliorer très significativement la qualité de la mesure. Aujourd'hui, on compte près de mille SNe dans le diagramme de Hubble, pour une précision sur w de l'ordre de 4%. Cette incertitude comprend à parts égales l'incertitude liée au nombre de SNe Ia utilisées (incertitude dite statistique) et les incertitudes liées aux imprécisions de l'analyse elle-même ainsi qu'à l'instrumentation (incertitude systématique). D'ici quelques années, le diagramme comptera plusieurs dizaines de milliers de SNe Ia, produites par les relevés ZTF, LSST et *Roman*. En combinaison avec les résultats de DESI, il sera possible de mesurer w avec une précision de l'ordre du pourcent, ainsi que d'éventuelles variations de w avec une précision de $\sim 10\%$.

Une difficulté centrale de l'analyse tient au fait que l'observateur mesure les flux de supernovæ redshiftées. Les flux mesurés dans les bandes passantes du télescope correspondent donc à des zones spectrales différentes d'une supernova à l'autre. Il est donc essentiel de pouvoir inférer à partir des courbes de lumière des SNe, les flux émis par une zone spectrale de référence (la même

pour toutes les SNe). Pour cela, on utilise un modèle spectrophotométrique, qui prédit l'évolution du spectre de chaque SN en fonction de la phase (i.e. la date rapportée au pic de luminosité). Un tel modèle est construit à partir des données de suivi spectroscopique et photométrique de SNe Ia. L'état de l'art est l'approche définie par le modèle SALT2 qui modélise le processus de mesure, ainsi que la diversité des SNe Ia.

SALT2 a été développé et entraîné sur des échantillons de quelques centaines de SNe. Il modélise les SNe Ia comme une famille à 2 paramètres. La croissance rapide des échantillons d'entraînement (quelques milliers de SNe aujourd'hui, quelques 10^4 d'ici 2025) justifie de revisiter la paramétrisation de SALT2. D'autres études, en particulier celles menées sur les séquences spectrophotométriques publiées dans le cadre du projet SNfactory suggèrent qu'un troisième paramètre pourrait améliorer la standardisation des SNe.

L'objet de ce travail de thèse a été le développement d'un successeur au cadre d'entraînement proposé par SALT2. Le code est écrit en `python` (avec très peu de `C` aux endroits critiques). Il a été désigné de façon à séparer la procédure d'ajustement du modèle, et rendre ainsi aisé le développement de paramétrisations alternatives à SALT2. Enfin, l'entraînement a été simplifié au maximum. Il est réduit à une unique minimisation, incluant le modèle et le modèle d'erreur associé. En particulier, la propagation des incertitudes de calibration photométriques est incluse dans l'entraînement du modèle. Le code est appelé `NaCl` pour Nouveaux Algorithmes de Courbes de Lumière.

Dans la première partie de ce document nous allons présenter le contexte général dans lequel s'inscrit la contrainte de l'équation d'état de l'énergie noire par la mesure de distances des SNe Ia. Le premier chapitre, introduit le modèle standard de la cosmologie. Le chapitre 2 décrit la physique des SNe Ia et leurs caractéristiques observationnelles, permettant leur standardisation. Le chapitre 3 est dédié à la construction du diagramme de Hubble des SNe Ia.

Dans une seconde partie, nous détaillons la construction de l'outil d'analyse développé durant cette thèse. Le chapitre 4 introduit, de manière chronologique la modélisation empirique des courbes de lumière SNe Ia. Le chapitre suivant décrit le modèle développé et utilisé, suivant une paramétrisation de la diversité des SNe Ia similaire au modèle SALT2. Le chapitre 6 détaille l'algorithme du minimiseur que nous avons développé pour ajuster simultanément le modèle moyen et sa variance, pris en compte dans un modèle d'erreur. Le chapitre 7 décrit les simulations que nous avons réalisées pour le développement des outils. Enfin les deux chapitres suivants caractérisent les biais et l'efficacité de l'estimateur développé.

Première partie

Caractérisation de l'énergie noire avec
les Supernovæ de type Ia

Le modèle standard de la Cosmologie

Dans cette partie, nous donnons quelques notions théoriques de la cosmologie moderne. Il s'agit de définir le formalisme, maintenant standard, fixant le cadre des analyses cosmologiques. Le modèle de l'Univers, s'est construit empiriquement, avec des découvertes parfois inattendues, comme l'accélération de l'expansion de l'Univers. Une première partie introduit la description de l'Univers par les équations d'Einstein. La partie suivante introduit la métrique décrivant la géométrie d'un Univers en expansion, homogène et isotrope aux échelles cosmologiques. Les différentes observables de distance sont définies dans la section suivante. La section 1.4 présente les équations de Friedmann, résolvant les équations d'Einstein grâce à la métrique définie précédemment. Le modèle d'Univers Λ CDM, est présenté section 1.5. Nous décrivons ensuite trois sondes de l'équation d'état de l'énergie noire. La section 1.7 détaille deux autres utilisations des supernovæ de type Ia pour contraindre d'autres paramètres cosmologiques. Le chapitre s'achève par les succès du modèle et les tensions qu'il reste à résoudre.

Contents

1.1	Les équations d'Einstein	15
1.2	Un univers isotrope, homogène et en expansion	15
1.2.1	Isotropie	15
1.2.2	Homogénéité	15
1.2.3	Loi de Hubble-Lemaître	15
1.2.4	Métrie FRLW	17
1.3	Observables	18
1.3.1	Décalage spectral ou <i>Redshift</i>	18
1.3.2	Les distances cosmologiques	19
1.4	Les équations de Friedmann	22
1.5	Le modèle ΛCDM	25
1.5.1	Le contenu énergétique de l'Univers	25
1.5.2	Le fond diffus cosmologique	32
1.5.3	Croissance des structures	33
1.6	L'étude de l'énergie noire avec la combinaison de différentes sondes cosmologiques	34
1.6.1	Les chandelles standards : SNe Ia	34
1.6.2	Le pic de l'oscillation acoustique des baryons	35
1.6.3	Combinaison de sondes	39
1.7	Autres utilisations cosmologiques des SNe Ia	41
1.7.1	Diagramme de Hubble pour mesurer H_0	41
1.7.2	Vitesses particulières	44
1.8	Succès & tensions du modèle	45

1.1 Les équations d'Einstein

La dynamique de l'Univers est décrite dans le cadre de la relativité générale par les équations d'Einstein [Einstein (1915)], qui lient la géométrie de l'espace-temps à la densité de ses différents constituants.

$$G_{\mu\nu} = -\frac{8\pi G}{c^4} T_{\mu\nu} \quad (1.1)$$

Le tenseur d'Einstein, $G_{\mu\nu}$ encode l'information géométrique de l'Univers. La partie droite de l'équation décrit son contenu énergétique : $T_{\mu\nu}$ est le tenseur énergie-impulsion, G est la constante universelle de gravitation et c est la vitesse de la lumière dans le vide. La résolution de ce système non-linéaire de 10 équations est complexe. Heureusement, l'introduction de certaines hypothèses, d'homogénéité et d'isotropie, simplifie leur résolution.

1.2 Un univers isotrope, homogène et en expansion

Pour résoudre les équations 1.1, on étend le principe copernicien à l'Univers tout entier, en supposant qu'il n'y existe pas d'observateur privilégié, donc qu'il est spatialement homogène et isotrope. Cette forte hypothèse, émise par Einstein et reprise par Lemaître et Friedmann, dans la première partie du XXème siècle, porte le nom de principe cosmologique. Ce principe, ajouté à l'observation d'un Univers en expansion permettent de définir une métrique décrivant simplement la géométrie de l'Univers.

1.2.1 Isotropie

L'isotropie de l'Univers est corroborée par l'observation au début des années 90 du fond diffus cosmologique (noté CMB pour *Cosmic Microwaves Background*) par le satellite COBE (*Cosmic Background Explorer*). La figure 1.1, présente la carte du ciel des faibles anisotropies en température du fond diffus cosmologique mesuré par le satellite PLANCK [Planck Collaboration et al. (2018)], preuve impressionnante de l'isotropie en température de l'Univers primordial.

1.2.2 Homogénéité

L'homogénéité de l'Univers est évidemment infirmée à l'échelle du système solaire ou des galaxies mais elle est confirmée aux échelles cosmologiques, à des distances de l'ordre du milliard d'années-lumière. La figure 1.2 représente plus de 200 000 galaxies observées par *2dF Galaxy Redshift Survey* [2dFGRS, Colless et al. (2001)] : l'homogénéité de l'Univers dépend manifestement de l'échelle considérée. Ainsi Hogg et al. (2005) ont montré que la répartition des galaxies lumineuses et rouges (LRG pour *Red Luminous Galaxies*) du relevé du *Sloan Digital Sky Survey* (SDSS) est compatible avec l'hypothèse d'homogénéité pour des échelles de $R \geq 70h^{-1}\text{Mpc}$.

1.2.3 Loi de Hubble-Lemaître

En 1927, Georges Lemaître [Lemaître (1927)] déduit des équations d'Einstein 1.1 que la vitesse d'éloignement entre deux objets est proportionnelle à leur distance. En 1929, Edwin Hubble fait la même déduction et relie empiriquement la vitesse d'éloignement, v_{gal} , de 32 galaxies à leur distance D_{gal} :

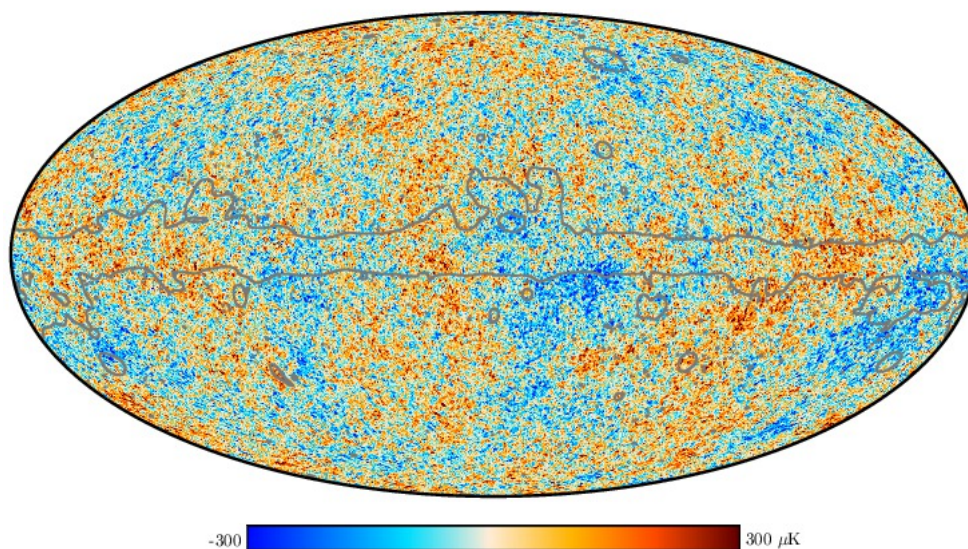


FIGURE 1.1 – Mesure des anisotropies de température du fond diffus cosmologique par le satellite PLANCK [Planck Collaboration et al. (2018)]. Les zones bleues correspondent aux régions "froides" et les rouges aux régions, "chaudes". La variation maximale entre deux régions de cette carte est de $600 \mu K$, autour de : $2.7255 K$ [Fixsen (2009)].

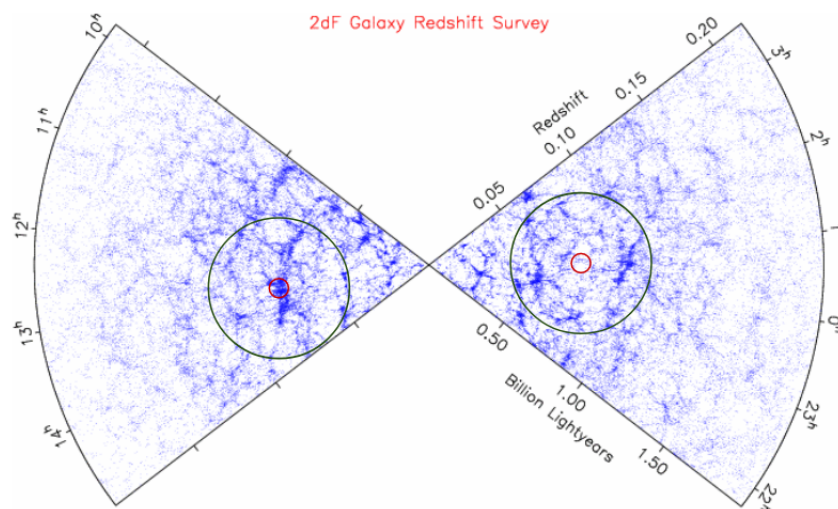


FIGURE 1.2 – Cartographie de plus de 200 000 galaxies réalisée par l'observatoire anglo-australien (AAO) sur une période de 5 ans dont la magnitude limite est de 19.5. On peut voir que deux régions "locales" de l'Univers, entourées en rouge, les densités de galaxies y sont différentes. Tandis que si l'on considère de plus grandes régions, en bleu, la densité de galaxies tend à être similaire.

$$v_{gal} = H_0 D_{gal} \quad (1.2)$$

La valeur de la constante de proportionnalité, H_0 , nommée constante de Hubble-Lemaître, mesurée par Hubble (1929), est de $500 \text{ km.s}^{-1}.\text{Mpc}^{-1}$ (cf. Figure 1.3).

Velocity-Distance Relation among Extra-Galactic Nebulae.

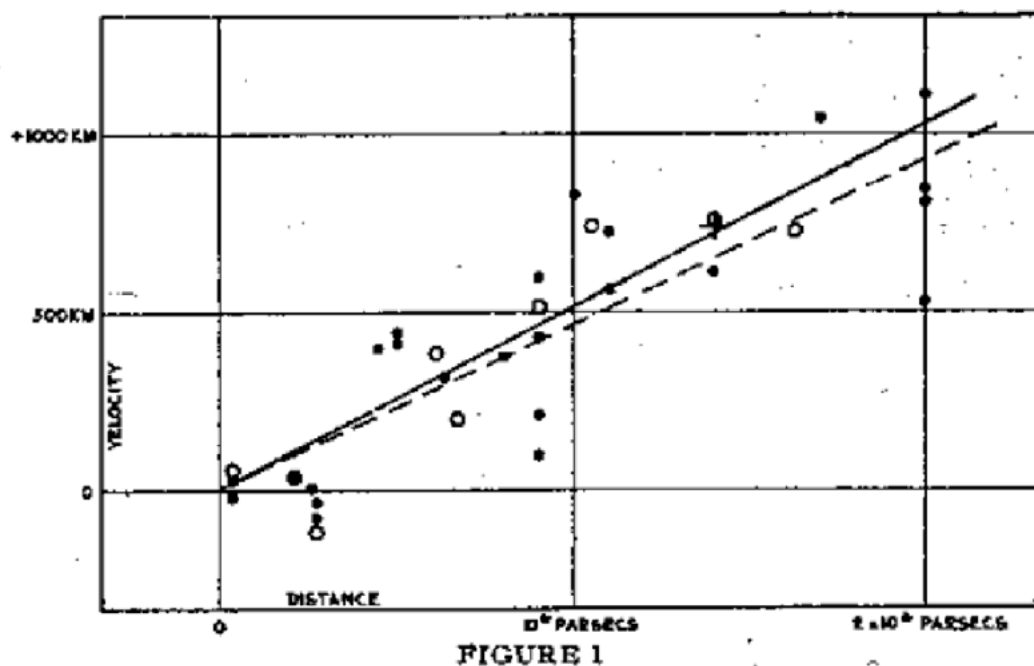


FIGURE 1.3 – Diagramme de Hubble [Hubble (1929)], vitesse d'éloignement de 32 galaxies en fonction de leur distance. Première observation de l'expansion de l'Univers.

À cette époque, la relation période luminosité, utilisée pour mesurer la distance de ces galaxies était mal calibrée, supposant les observations de chandelles standards¹ à grande distance, alors que ces observations étaient en réalité des clusters d'étoiles et non des étoiles individuelles. La mesure du paramètre H_0 a ensuite convergé vers une valeur proche de $70 \text{ km.s}^{-1}.\text{Mpc}^{-1}$, grâce à l'accord de différentes sondes. Aujourd'hui, la mesure précise de H_0 et de son évolution sont de grands enjeux de la cosmologie moderne sur lesquels nous reviendrons dans la suite de ce chapitre. Cependant le paramètre H a évolué au cours de l'histoire de l'Univers. La mesure de $H(t)$ est une mesure de l'histoire de l'expansion de l'Univers au temps t . Du fait de la proximité des galaxies observées par Hubble, cette constante est mesurée au temps actuel de l'Univers, $t = t_0$, noté $H_0 = H(t = t_0)$.

1.2.4 Métrique FRLW

La géométrie de l'espace-temps est décrite par le tenseur métrique, terme de droite de l'équation 1.1, $g_{\mu\nu}$, qui lie les distances physiques des intervalles d'espace-temps ds aux coordonnées et définie par la formule suivante :

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (1.3)$$

Pour un Univers en expansion, homogène et isotrope, la métrique dite de Friedmann-Robertson-Lemaître-Walker (FRLW) [Weinberg (1972)] s'écrit simplement en coordonnées polaires dans le

1. Object de luminosité connue.

référentiel de l'observateur (t, r, θ, ϕ) :

$$ds^2 = c^2 dt^2 - a(t) \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta + \sin^2(\theta)d\phi^2) \right] \quad (1.4)$$

Le facteur d'échelle $a(t)$ décrit une contraction ou expansion de l'espace, ce qui implique que la distance entre deux objets varie au cours du temps. Le système de coordonnées (r, θ, ϕ) est dit *comobile*, c'est-à-dire que les coordonnées d'un objet fixé ne varient pas avec l'expansion de l'Univers. k est la *courbure de l'espace*, qui définit la géométrie des sections spatiales de l'espace-temps. $k = -1, 1$ ou 0 pour un Univers respectivement fermé ou sphérique, ouvert ou hyperbolique, plat. Il existe différents systèmes de coordonnées plus ou moins pratiques. Par convention, on réalise le changement de variable $d\chi = \frac{dr}{\sqrt{1-kr^2}}$, puisque la partie radiale est la seule à dépendre de k et on obtient donc :

$$r = S_k(\chi) \equiv \begin{cases} \sin(\chi) & \text{si } k = -1 \\ \sinh(\chi) & \text{si } k = 1 \\ \chi & \text{si } k = 0 \end{cases}$$

On peut ré-écrire, la métrique FRLW dans le système de coordonnées (χ, θ, ϕ) :

$$ds^2 = c^2 dt^2 - a(t) [d\chi^2 + S_k(\chi)(d\theta + \sin^2(\theta)d\phi^2)] \quad (1.5)$$

Ainsi dans ce cadre, la métrique de l'Univers est simplement décrite par le paramètre k et la loi d'évolution du facteur d'échelle $a(t)$. Le taux d'expansion de l'Univers s'exprime alors comme :

$$H(t) = \frac{\dot{a}}{a} \quad (1.6)$$

1.3 Observables

En pratique nous avons besoin d'observables, tel que le décalage spectral ou des distances, permettant de contraindre l'histoire de l'expansion de l'Univers, $a(t)$.

1.3.1 Décalage spectral ou *Redshift*

L'expansion de l'Univers implique que deux points distants aux échelles cosmologiques s'éloignent l'un de l'autre. Un photon émis par une galaxie à une longueur d'onde λ_e et reçu sur Terre sera affecté par un effet similaire à un effet Doppler-Fizeau dû à la vitesse d'éloignement de la source du fait de l'expansion de l'Univers. Il sera reçu à une longueur d'onde λ_o , décalée vers le rouge. Ce décalage spectral est nommé *redshift*. Le photon, émis au temps t_e et reçu à t_o , suit une géodésique pour laquelle $ds^2 = 0$ ce qui implique $d\chi = \frac{cdt}{a(t)}$. Considérons maintenant deux photons, émis respectivement à t_e et $t_e + \delta t_e$ par une source à des coordonnées comobiles $(\chi, 0, 0)$ et observés à t_o et $t_o + \delta t_o$. On obtient, en considérant que pendant la période d'émission de l'onde le facteur d'échelle est constant :

$$\chi = \int_{t_e + \delta t_e}^{t_o + \delta t_o} \frac{cdt}{a(t)} = \int_{t_e}^{t_o} \frac{cdt}{a(t)} \Rightarrow \int_{t_e + \delta t_e}^{t_e} \frac{cdt}{a(t)} = \int_{t_o + \delta t_o}^{t_o} \frac{cdt}{a(t)} \Rightarrow \frac{\delta t_e}{a(t_e)} = \frac{\delta t_o}{a(t_o)}$$

Or, $\delta t_X \propto \lambda_X$ qui implique $\frac{\lambda_e}{a(t_e)} = \frac{\lambda_o}{a(t_o)}$. On définit le *redshift* comme le rapport des longueurs d'onde émise et reçue :

$$1 + z = \frac{\lambda_o}{\lambda_e} = \frac{a(t_o)}{a(t_e)} \quad (1.7)$$

Le *redshift* est donc une mesure directe du rapport du facteur d'échelle entre l'émission et la réception du photon.

1.3.2 Les distances cosmologiques

Il existe différents types de distances aux échelles cosmologiques, égales à petit *redshift*, qu'il convient de bien définir pour comprendre leur distinction à grand *redshift*.

Distance physique La distance la plus simple à laquelle on puisse penser, étant donnée la métrique utilisée dans équation 1.5, est la distance physique. Elle est définie à t_0 entre deux points $(0, \theta, \phi)$ et (χ, θ, ϕ) par :

$$d_p(z) = a_0 S_k(\chi) \quad (1.8)$$

Pour mesurer, χ , il est pratique de l'exprimer en fonction du *redshift*, seule observable directe, et des paramètres cosmologiques. En considérant un photon émis en $(\chi, 0, 0, t)$ et reçu à $(0, 0, 0, t_0)$, on obtient la formule suivante :

$$\chi = c \int_t^{t_0} \frac{dt}{a(t)} \quad (1.9)$$

On procède ensuite aux changements de variables $dt = da/\dot{a} = \frac{da}{Ha}$ puis $\frac{dz}{da} = \frac{1}{a^2}$, pour obtenir finalement :

$$\chi = c \int_0^z \frac{d\tilde{z}}{H(\tilde{z})} \quad (1.10)$$

Des équations 1.8 et 1.10, on déduit la formule de la distance comobile en fonction du *redshift* :

$$d_p(z) = a_0 S_k \left(\frac{c}{a_0} \int_0^z \frac{d\tilde{z}}{H(\tilde{z})} \right) \quad (1.11)$$

Cette distance n'est en pratique pas observable directement, puisqu'il faudrait pouvoir observer la valeur de H à tous les *redshift* sur la ligne de visée, pour en calculer l'intégrale. Par contre, il existe d'autres distances observables, desquelles on peut inférer des contraintes sur les paramètres cosmologiques.

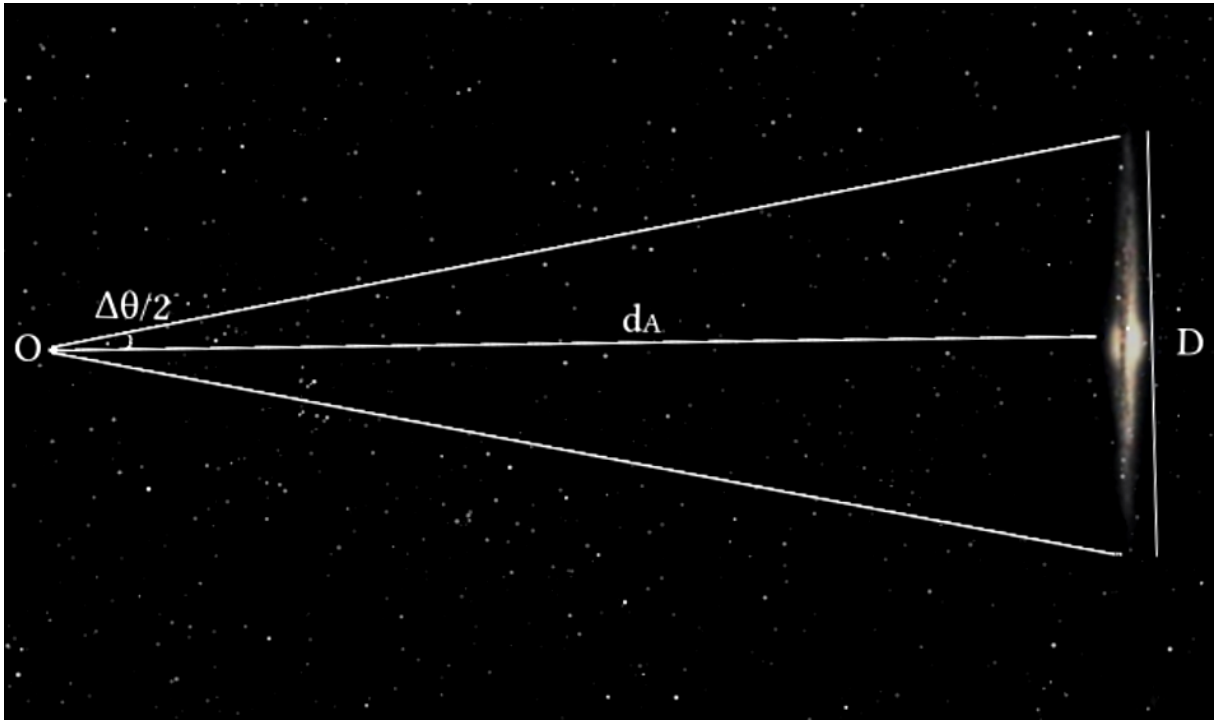


FIGURE 1.4 – Distance angulaire rapport de la taille d’un objet D et de son angle apparent $\Delta\theta$.

Distance angulaire En pratique, une distance mesurable est celle qui lie la taille d’un objet lointain à son diamètre apparent. La distance angulaire d’un objet d_A est définie par le rapport de sa taille D et de son angle apparent (cf. Figure 1.4). Cette distance est une fonction de la coordonnée comobile χ :

$$d_A = \frac{D}{\Delta\theta} = a(t)S_k(\chi) \quad (1.12)$$

Soit, en utilisant les équations 1.7 et 1.10 :

$$d_A(z) = \frac{1}{1+z} a_0 S_k \left(\frac{c}{a_0} \int_0^z \frac{d\tilde{z}}{H(\tilde{z})} \right) \quad (1.13)$$

Ainsi, en mesurant le *redshift* d’un objet, sa taille et son angle apparent, on peut déduire des contraintes sur le taux d’expansion de l’Univers et sa courbure.

Distance de luminosité La distance de luminosité (cf. figure 1.5) lie la luminosité d’une source, L , au flux surfacique reçu par l’observateur, f (énergie par unité de temps et de surface) par la formule suivante :

$$d_L^2 = \frac{L}{4\pi f} \quad (1.14)$$

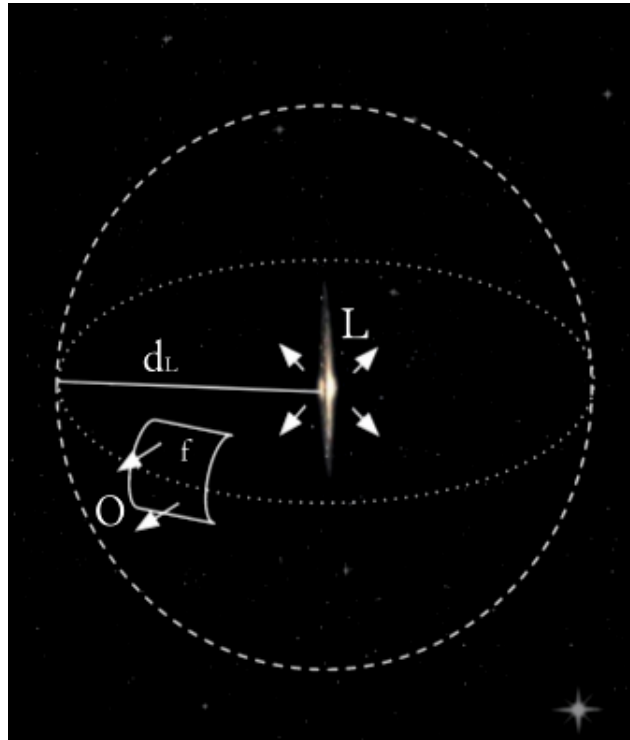


FIGURE 1.5 – Distance de luminosité d_L d'une source de luminosité intrinsèque L . L'observateur reçoit un flux surfacique.

Les photons émis à t par la source située aux coordonnées $(\chi, 0, 0)$, se propagent jusqu'à l'observateur situé en $(0, 0, 0)$ et sont répartis sur une sphère de centre χ , de rayon $a_0 S_k(z)$ et de surface $S = 4\pi a_0^2 S_k^2(z)$.

Comme on l'a vu précédemment, lorsque qu'un photon traverse des distances cosmologiques, sa longueur d'onde est décalée vers le rouge, donc son énergie, $\frac{hc}{\lambda}$, est divisée par un facteur $1 + z$. De plus, si on considère des photons émis pendant un intervalle de temps dans le référentiel de la source, alors ils seront reçus dans un intervalle de temps dilaté d'un même facteur $1 + z$ dans le référentiel de l'observateur.

Le flux reçu par l'observateur est relié à la luminosité L de la source par : $f = \frac{L}{S(1+z)^2}$ On en déduit :

$$f = \frac{L}{4\pi a_0^2 S_k^2(\chi)(1+z)^2} \quad (1.15)$$

On obtient donc par identification de 1.14 et 1.15 et en utilisant de l'équation 1.10 :

$$d_L(z) = a_0 S_k \left(\frac{c}{a_0} \int_0^z \frac{d\tilde{z}}{H(\tilde{z})} \right) (1+z) \quad (1.16)$$

On voit que ces différentes distances sont liées seulement par des facteurs $1 + z$. Elles sont asymptotiquement égales pour $z \ll 1$ et divergentes à grand *redshift* (cf. Figure 1.6) :

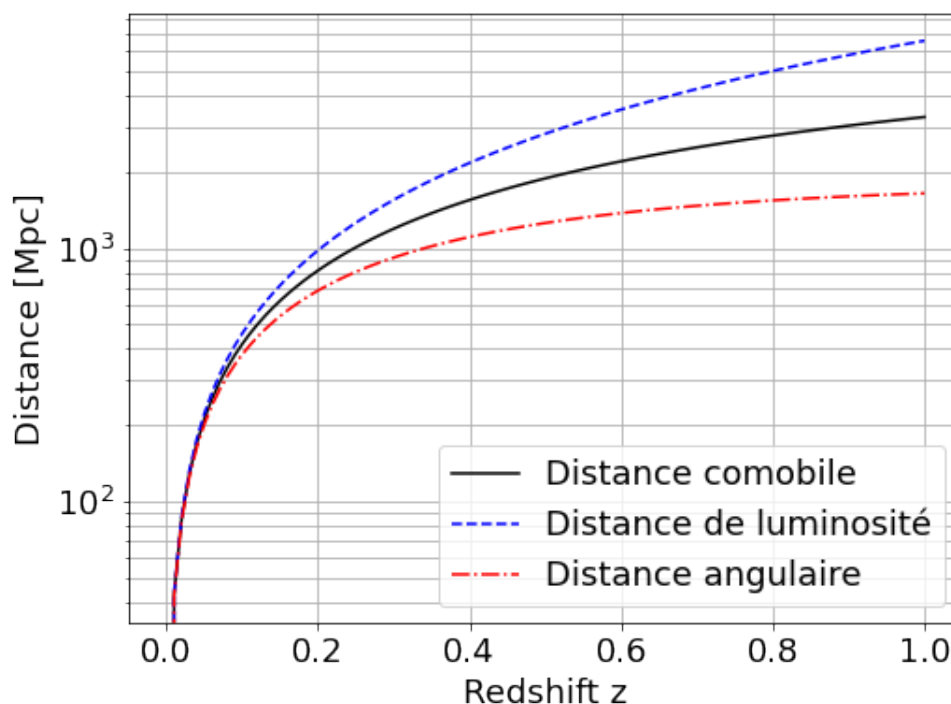


FIGURE 1.6 – Comparaison des différentes distances (calculées pour un même jeu de paramètres cosmologiques) en fonction du *redshift*. À bas *redshift*, elles sont asymptotiquement égales.

$$d_L(z) = d_p(z)(1 + z) = d_A(z)(1 + z)^2 \quad (1.17)$$

À petite distance, donc à petit *redshift*, on retrouve la loi linéaire de Hubble :

$$d_L(z) \sim d_A(z) \sim \frac{cz}{H_0} \quad (1.18)$$

1.4 Les équations de Friedmann

Maintenant que nous avons une description géométrique de l'Univers, les équations d'Einstein peuvent être résolues pour obtenir une expression de $a(t)$. Le tenseur d'Einstein, introduit dans l'équation 1.1, s'écrit :

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \quad (1.19)$$

$R_{\mu\nu}$ est le tenseur de Ricci, la contraction du tenseur de courbure de Riemann et R est sa trace. On peut montrer, avec la métrique FRLW (équation 1.5) que le tenseur de Ricci s'écrit :

$$R_{\mu\nu} = a(t)^2 \left(\frac{\ddot{a}(t)}{a(t)} + 2\frac{\dot{a}^2(t)}{a^2(t)} + 2\frac{k}{a^2(t)} \right) g_{\mu\nu} \quad (1.20)$$

En supposant que le contenu énergétique de l'Univers est décrit par un ou plusieurs fluides parfaits à l'équilibre, dont la pression p et la densité ρ dépendent uniquement du temps puisque l'Univers est homogène et isotrope, le tenseur énergie-impulsion $T_{\mu\nu}$ s'écrit :

$$T_{\mu\nu} = \begin{pmatrix} -\rho(t) & 0 & 0 & 0 \\ 0 & p(t) & 0 & 0 \\ 0 & 0 & p(t) & 0 \\ 0 & 0 & 0 & p(t) \end{pmatrix}$$

En 1924, Friedman en déduit les équations différentielles dites de Friedman-Lemaître, qui lient la dynamique de l'expansion de l'Univers donc $a(t)$ à son contenu énergétique :

$$H^2(t) = \left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{kc^2}{a^2(t)} \quad (1.21)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \left(\rho + 3\frac{p}{c^2} \right) \quad (1.22)$$

En dérivant la première équation de Friedmann-Lemaître et en identifiant avec le terme de droite de la seconde, on obtient l'équation de conservation de l'énergie :

$$\dot{\rho} + 3\frac{\dot{a}}{a}(p + \rho) = 0 \quad (1.23)$$

On peut simplifier cette équation en utilisant l'équation d'état des gaz parfaits, liant la pression p à la densité d'un gaz ρ : $p = w\rho$. On obtient en intégrant l'équation 1.23, l'évolution de la densité des différentes composantes de l'Univers en fonction du paramètre w et du facteur d'échelle :

$$\rho(t) = \rho_0 \left(\frac{a}{a_0} \right)^{-3(1+w)} \quad (1.24)$$

On peut ainsi exprimer l'évolution du facteur d'échelle en fonction du temps, lorsque la dynamique de l'Univers est dominée par une seule composante de paramètre w :

$$a(t) \propto \begin{cases} t^{\frac{2}{3(1+w)}} & w \neq -1 \\ e^{Ht} & \text{sinon.} \end{cases} \quad (1.25)$$

Ceci implique que l'Univers est en expansion accélérée si et seulement si :

$$\ddot{a} > 0 \Leftrightarrow w < -1/3 \quad (1.26)$$

La division la plus simple pour le contenu de l'Univers est la distinction entre particules relativistes, désignées par un indice r , et particules non relativistes, désignées par un indice m . On regroupe les particules par classe ayant la même équation d'état. Pour un Univers dominé par la matière non-relativiste, nous savons que la pression d'un gaz est donnée par l'agitation thermique de ses atomes ou molécules. La vitesse d'agitation thermique est proche de celle du son $c_s = 300 \text{ m.s}^{-1}$. Pour un tel gaz, sa pression est $P \propto \rho v^2$, où v est la vitesse d'agitation thermique. Cette pression est largement inférieure à ρc^2 , ce qui est gravitationnellement négligeable. Le paramètre d'état de cette matière est donc nul : $w_m = v^2/c^2 \sim 0$, impliquant

$$\rho_m \propto a^{-3} \propto (1+z)^3 \quad (1.27)$$

Pour la particule relativiste, le facteur w_r est donné par la relation de pression : $p_r = \frac{1}{3}\rho_r c^2$, on a donc $w_r = \frac{1}{3}$ ce qui implique :

$$\rho_r \propto a^{-4} \propto (1+z)^4 \quad (1.28)$$

On peut exprimer ces différentes densités des composantes de l'Univers en unités de *densité critique*, ρ_c , définies comme la densité d'un Univers de courbure spatiale nulle ($k = 0$) :

$$\rho_c = \frac{3H_0^2}{8\pi G} \quad (1.29)$$

On définit alors pour chaque composante, X , les densités aux valeurs actuelles ($t = t_0$) en unités de densité d'énergie critique :

$$\Omega_X = \frac{\rho_X}{\rho_c} \quad (1.30)$$

On peut maintenant, ré-exprimer l'équation 1.21, en fonction des Ω_X et du *redshift* z :

$$H^2(z) = H_0^2[\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_k(1+z)^2] \quad (1.31)$$

où on définit $\Omega_k = \frac{-kc^2}{R_0^2 H_0^2}$. Pour $z = 0$, on a :

$$\sum_X \Omega_X = \Omega_m + \Omega_k + \Omega_r = 1 \quad (1.32)$$

1.5 Le modèle Λ CDM

Le modèle Λ CDM est un modèle décrit par 6 paramètres libres. Ce nombre de paramètres est faible comparé au nombre d'observations qui le valident. Par comparaison, le modèle standard de la physique des particules possède 18 paramètres libres. Ces six paramètres sont :

- le paramètre de densité de matière noire actuelle, Ω_{cdm} .
- le paramètre de densité de matière baryonique actuelle, Ω_b
- la constante de Hubble mesurée aujourd'hui, H_0 .
- l'épaisseur optique de réionisation τ , qui quantifie le nombre moyen de diffusions Thomson subies par un photon du fond diffus cosmologique lors de la réionisation.
- l'indice et l'amplitude des perturbations scalaires, n_s et A_s , qui définissent le spectre de puissance des fluctuations primordiales : $P_s(k) = A_s k^{n_s-1}$.

La définition et la mesure des quatre premiers paramètres sont présentées dans cette section.

1.5.1 Le contenu énergétique de l'Univers

Pour reconstruire l'histoire de l'expansion de l'Univers (cf.équation 1.31), il faut paramétriser son contenu énergétique.

Particules non-relativistes

Matière baryonique Le terme de *matière baryonique* comprend toute la matière visible non-relativiste. Ce sont les protons et les neutrons, qui sont des *baryons*² qui contribuent principalement à la densité d'énergie. Lorsque l'on suppose que l'Univers est uniquement composé de matière visible, i.e. qui interagit par interaction électromagnétique, il apparaît des problèmes entre théorie et observations. C'est pourquoi une matière interagissant seulement par gravité à été introduite : la matière noire.

Matière noire - Preuves indirectes En 1937, Fritz Zwicky [Zwicky (1937)] essaie de déduire la masse totale de l'amas de Coma (Abell 1656) en calculant la dispersion des vitesses des galaxies qui le constituent. Il compare cette masse dynamique à la masse *lumineuse*, calculée à partir de la quantité de lumière émise par l'amas et en déduit un défaut de masse impliquant l'existence de matière invisible. Cette méthode se base sur deux hypothèses difficiles à justifier. La première est de supposer que le système est en équilibre stationnaire, dit *virialisé* et la seconde est l'isotropie de la distribution de vitesse.

En 1970, cette première détection indirecte de matière noire est confirmée par Vera Rubin [Rubin & Ford (1970)], en étudiant la courbe de rotation, distribution de vitesse orbitale en fonction de la distance au centre, de la galaxie spirale d'Andromède (M31). La vitesse orbitale est mesurée par effet Doppler dans différentes longueurs d'onde, en radio, à partir de la raie à 21 cm de l'hydrogène neutre HI et en optique, à partir du spectre d'émission de l'hydrogène chargé HII. L'hydrogène étant la principale composante du gaz de la galaxie, on peut donc sonder la distribution de masse au-delà de sa partie visible. Considérons une particule de masse m , située à une distance r du centre d'une galaxie de profil de masse $M(< r)$. Supposons que cette particule possède une vitesse v dans le référentiel de la galaxie, alors d'après le principe fondamental de la dynamique :

$$v = \sqrt{\frac{GM(< r)}{r}}$$

2. Particules composées de trois fermions appelés quarks. En cosmologie, le terme de matière Baryonique renvoie à la famille des particules non-relativistes, dont la majorité sont des baryons.

Les observations (cf. Figure 1.7) montrent un profil de vitesse plat et donc qu'une grande quantité de matière s'étend loin du centre galactique, puisque $M(< r) \propto r$. On peut, là aussi, comparer cette distribution de masse dynamique estimée à partir de la distribution de luminosité de la galaxie, et constater que la platitude de la courbe de rotation à grand r est due à un halo de matière invisible, dite *noire*. Vera Rubin introduit le terme de matière sombre et montra dans des études ultérieures que ce phénomène n'est pas marginal mais commun à toutes les galaxies.

Recherche par mesure directe - [Mayet et al. (2016)] - Il existe d'autres preuves indirectes de l'existence de la matière noire mais aucune détection directe n'a été réalisée. Nous n'avons que très peu d'informations sur les particules de matière noire, si ce n'est que ce sont des particules qui interagissent au moins par gravitation avec la matière visible mais pas par électro-magnétisme. De plus, pour les modèles cosmologiques les plus éprouvés la matière noire serait dite *froide* (CDM, en anglais pour Cold Dark Matter), i.e. les particules se déplacent lentement par rapport à la vitesse de la lumière. Il existe différents modèles de matière noire froide parmi lesquels on peut citer les deux plus populaires :

- les WIMPs (*Weakly Interacting Massive Particules*) sont des particules appartenant à une famille de modèles étendus du Modèle Standard de la physique des particules, les modèles dit de Supersymétrie. Ces modèles prédisent de bons candidats pour la matière noire froide mais aucune des particules prédites par ces modèles n'a pour l'instant été observée...
- les axions sont des particules introduites dans le but d'empêcher la violation de la symétrie charge parité de l'interaction forte. Cette symétrie implique que la physique soit la même pour une particule ou une antiparticule, quelles que soient leurs positions dans l'espace. Ces particules légères et froides auraient pu être produites dans l'Univers primordial et seraient toujours présentes aujourd'hui dans les halos de galaxies.

Plusieurs expériences de détection directe sont en cours actuellement. Sur la figure 1.8, sont présentées les limites supérieures, sur la probabilité d'interaction de ces particules en fonction de leur masse, imposées par quelques unes de ces expériences. En effet, les différents modèles de matière noire prédisent des particules de section efficace d'interaction, variant grandement en fonction de la masse (cf. Figure 1.8). Les limites d'exclusions décroissent avec la masse des particules puisque la sensibilité du détecteur, qui est une caractérisation du bruit de fond à partir d'un certain seuil, dépend de la masse de ces particules. Ainsi les particules les plus lourdes avec des probabilités d'interaction plus élevés sont plus efficacement exclues.

Certaines expériences sont en cours comme DAMIC [Aguilar-Arevalo et al. (2015)], ou encore en construction XENONnT [Aprile et al. (2017)].

Actuellement, les mesures les plus précises de densité de matière baryonique et noires dans l'Univers sont données par l'expérience Planck [Planck Collaboration et al. (2020)], dont les valeurs sont respectivement 4.9% et 26.4% du budget énergétique totale de l'Univers.

Particules relativistes

Il y a principalement deux types de particules relativistes qui nous intéressent. Le **photon**, médiateur de l'interaction électromagnétique, est le principal vecteur d'information que nous possédons entre les différents objets composant l'Univers et nous. Les **neutrinos**, notés ν , sont une classe de particules appartenant aux fermions, ayant une charge leptonique et n'interagissant qu'avec les autres leptons, par interaction faible. Ces particules ont été introduites par Pauli en 1930 pour expliquer le spectre continu de la désintégration β . La première observation a été réalisée en 1956, par J.L. Cowan et F. Reines. Le modèle standard de la physique des particules ne prédit pas de masse aux neutrinos. Pourtant les observations de l'expérience SuperKamiokandé

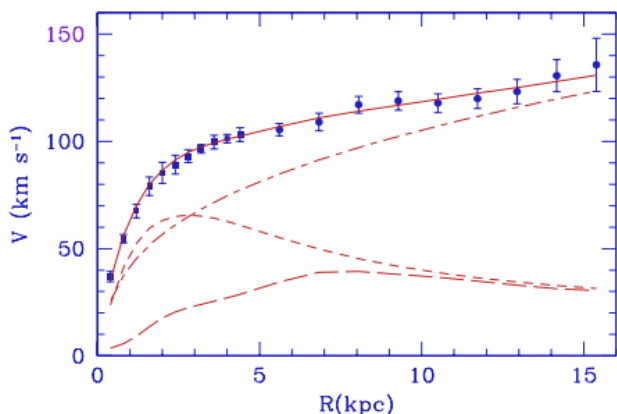


FIGURE 1.7 – Courbe de rotation de la galaxie M33 (points bleus). En rouge la somme des différentes contribution du modèle, en tiret-point la contribution du halo de matière noire, en tiret court la contribution du disque stellaire et en tiret long la contribution du gaz interstellaire [Corbelli & Salucci (1999)].

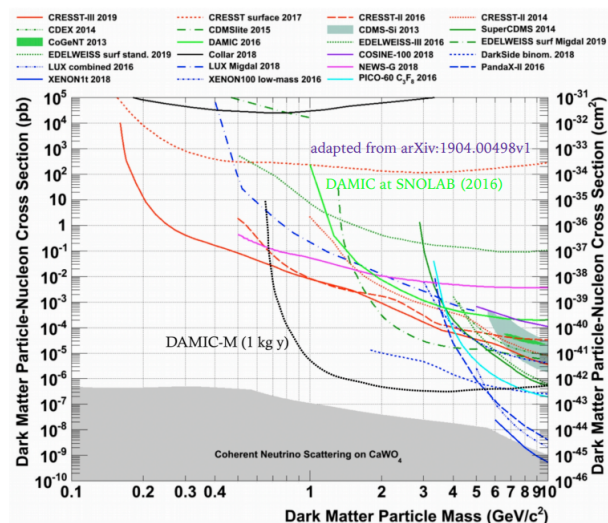


FIGURE 1.8 – Graphique des limites supérieures de sections efficaces d’interaction en fonction de la masse des particules de matière noire, établies par les différentes expériences.

[Fukuda et al. (1998)] ont mis en évidence dès 1998, un mécanisme permettant aux neutrinos de changer de saveur, i.e. de charge leptonique (électronique, muonique ou tauique) pendant leur propagation, appelé *oscillations des neutrinos*. Ce mécanisme prédit des masses non nulles pour les neutrinos. L’étude de la formation des grandes structures de l’Univers est une sonde pour contraindre la somme des masses des neutrinos.

La contribution actuelle (à t_0) de ces deux composantes au contenu énergétique de l’Univers est négligeable $\Omega_r \sim 8 \times 10^{-5}$.

Énergie noire

La magnitude apparente d’une source lumineuse est définie comme la mesure de son flux, f rapporté au flux connu d’une étoile dite standard f_0 en unités logarithmiques :

$$m = -2,5 \log \left(\frac{f}{f_0} \right) \quad (1.33)$$

Elle est reliée à la luminosité intrinsèque L de l’objet et sa distance de luminosité d_L (cf. équation 1.14), par la relation suivante :

$$m = -2,5 \log \left(\frac{L}{f_0 4\pi d_L^2} \right) \quad (1.34)$$

On peut introduire M , la magnitude absolue de la source, correspondant à la magnitude apparente de l’objet s’il était situé à une distance de 10 pc :

$$M = -2,5 \log \left(\frac{L}{f_0 4\pi (10pc)^2} \right) \quad (1.35)$$

On a alors la relation : $m = M + \mu$ ou le module de distance, μ est défini comme :

$$\mu = 5 \log \left(\frac{d_L}{10pc} \right) \quad (1.36)$$

Sur la figure 1.9, on retrouve le diagramme de Hubble de la figure 1.3. Les courbes bleues représentent différents modèles cosmologiques pour lesquels, Ω_m varie de 0.1 à 2.5 et Ω_r est négligé. On remarque qu'il était à l'époque de Hubble impossible de mettre une contrainte sur la densité des fluides, puisque à bas *redshift* toutes les courbes bleues se superposent.

La solution pour pouvoir discriminer parmi les différents modèles est d'ajouter des mesures à plus haut *redshift* dans la région où les courbes divergent. Ceci a été fait, à la fin des années 1990, jusqu'à $z \sim 0.5$ en utilisant comme indicateur de distance, non plus des galaxies mais des supernovæ de type Ia (SN Ia) [Riess et al. (1999), Perlmutter et al. (1999)]. Sur la figure 1.10, sont présentées les données d'une extension du diagramme de Hubble des SN Ia à plus grand *redshift*, lors d'une analyse nommée JLA pour *Joint Ligth curve Analysis* [Betoule et al. (2014)]. Les lignes bleues sont toujours des modèles d'Univers dominés par la matière avec différentes densités des matières, de 0.1 à 2.5 : aucune d'elles ne décrit le diagramme de Hubble des supernovæ. La seule présence de matière ne permet pas de décrire les observations. C'est la première mise en évidence de l'accélération de l'expansion de l'Univers.

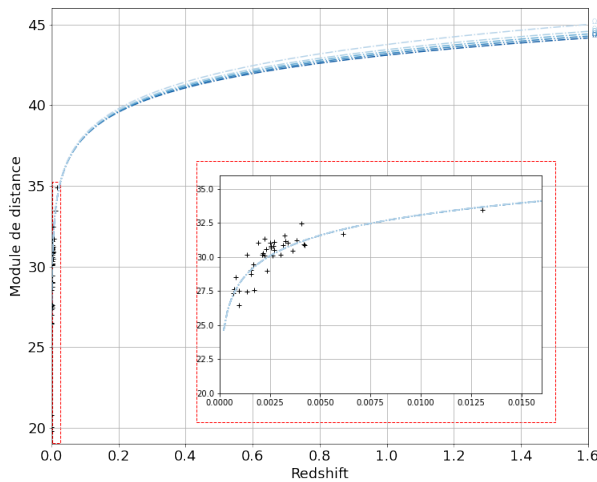


FIGURE 1.9 – Diagramme de Hubble avec les galaxies observées par Hubble [Hubble (1929)], jusqu'au *redshift* de 0.015. Les courbes bleues représentent la prédiction théorique de modèles cosmologiques contenant seulement de la matière et des particules relativistes.

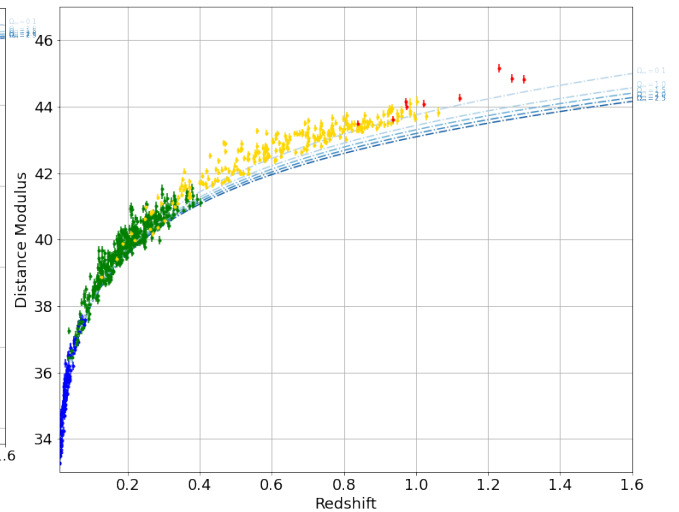


FIGURE 1.10 – Diagramme de Hubble étendu à partir de donnée de JLA. Les courbes théoriques d'Univers composées seulement de matière ne décrivent pas les observations.

En effet une solution pour obtenir une meilleure description des données est d'ajouter un nouveau fluide capable d'expliquer cette accélération de l'expansion de l'Univers, donc de para-

mètre d'équation d'état $w < -1/3$, l'énergie noire, noté EN , de densité critique Ω_{EN} . Sur la figure 1.11, ont été ajoutés en rouge des modèles d'Univers plats – c'est à dire pour lesquels $\Omega_k = 0 \Rightarrow \sum_X \Omega_X = 1$ – composés uniquement de matière Ω_m et d'énergie noire Ω_{EN} . On a donc $\Omega_{EN} + \Omega_m = 1$. Ainsi le modèle décrivant le mieux les données correspond à un Univers

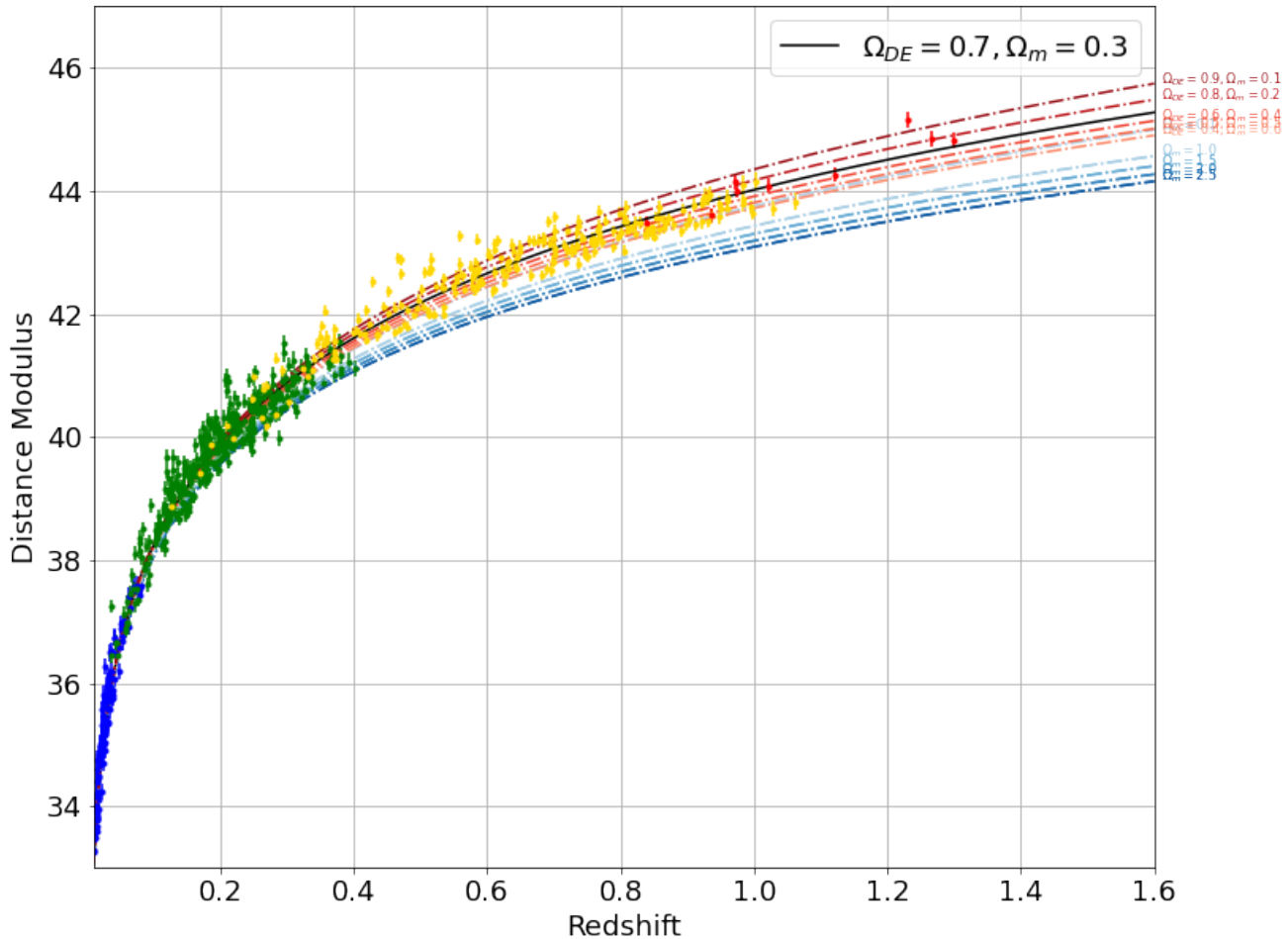


FIGURE 1.11 – Diagramme de Hubble étendu grâce aux données de JLA avec les prédictions de modèles d'Univers composés de matière, courbes bleues et de modèle d'Univers plats composés de matière et d'énergie noire (sous forme de constante cosmologique), en rouge. Le modèle décrivant le mieux les données est composé à environ 70% d'énergie noire et 30 % de matière (baryonique et noire).

composé à 30% de matière et 70% d'énergie noire dans l'Univers [Scolnic et al. (2018)]. Un tel Univers se retrouve donc en expansion accélérée.

Constante Cosmologique Historiquement, Einstein a introduit dans l'équation 1.1 une constante, Λ pour obtenir comme solution un Univers statique [Einstein (1917)] :

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = -\frac{8\pi G}{c^4}T_{\mu\nu} \quad (1.37)$$

Dès l'observation de l'expansion de l'Univers par Hubble, cette constante fut retirée des équations. Cependant, ce terme est réapparu dans les équations car il permet d'obtenir une expansion accélérée de l'Univers, avec une paramétrisation simple de l'énergie noire. Il y a deux manières de considérer cette constante cosmologique dans l'équation 1.37, mathématiquement équivalentes mais physiquement très différentes :

- Si on place le nouveau terme à gauche, alors nous avons une modification de la géométrie de l'Univers, donc de la relativité générale.
- Par contre si on le place à droite, on change le contenu énergétique de l'Univers, en ajoutant un nouveau fluide, l'*énergie noire*, avec une équation d'état, $w = -1$. Cette valeur implique une densité constante en fonction du facteur d'échelle (équation 1.24). Cette énergie noire est appelée constante cosmologique.

Il existe des modèles de gravité modifiée qui tentent d'expliquer l'accélération de l'expansion de l'Univers sans ajout d'énergie noire. Le modèle standard de la cosmologie Λ CDM est principalement composé de matière noire froide (*CDM* pour *Cold Dark Matter*) et d'énergie noire, analogue à une constante cosmologique.

Un nouveau fluide On introduit donc un nouveau fluide dans le contenu de l'Univers, d'équation d'état : $p_{EN} = w\rho_{EN}$, avec $w < -1/3$.

Puisque nous ne connaissons pas la nature de l'énergie noire, il est courant de proposer des extensions du modèle Λ CDM. Une paramétrisation tenant compte d'une possible variation du paramètre d'état avec le *redshift* peut être définie de la manière suivante :

$$w(z) = w_0 + w_a \frac{1}{1+z} \quad (1.38)$$

Dans le cadre de Λ CDM, nous avons $w_0 = -1$ et $w_a = 0$. Nous pouvons ainsi réécrire l'équation de Friedmann 1.21 :

$$H^2(z) = H_0^2 \left[\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_k(1+z)^2 + \Omega_{EN}(1+z)^{3(1+w_0+w_a)} e^{(-3w_a \frac{z}{1+z})} \right] \quad (1.39)$$

On retrouve bien le fait que lorsque w est constant et égal à -1, la densité d'énergie noire est bien une constante cosmologique et ne varie pas avec le *redshift*. Lorsque que l'on parle de constante cosmologique cela implique que l'on suppose $w = -1$. Par contre lorsque l'on parle d'énergie noire, sa valeur peut être différente. Il est à noter que l'énergie noire est responsable d'une force répulsive, par opposition aux forces engendrées par les différentes autres quantités d'énergie et que dans le cas d'une constante cosmologique (valeur conformes pour toutes les observations de w_{DE}

actuelle) sa densité d'énergie est constante malgré l'expansion de l'Univers. Cette dynamique est très différente de l'évolution de la densité de matière (cf. Figure 1.12).

La figure 1.12 représente les différentes densités d'énergie en fonction de l'inverse de $1 + z$. L'évolution des densités de radiation, matière et énergie noire avec le *redshift* sont représentées respectivement en rouge, bleu et vert. Lorsque l'on considère un Univers plat, $\Omega_k = 0$, on distingue trois phases dans l'histoire de l'Univers, représentées par les lignes grises et verticales. Une première phase pour laquelle la densité de particules relativistes dominait, puis une seconde où la matière dominait, et enfin récemment une phase pour laquelle l'énergie noire domine. La valeur de la courbe verte dépend de l'équation d'état de l'énergie noire, w_0 et w_a et la zone verte représente une zone pour laquelle l'énergie noire s'écarte de la constante cosmologique. La mesure de ces paramètres d'état est un enjeu important pour comprendre l'évolution de l'histoire de l'Univers et la nature de l'énergie noire. La mesure de distance dans l'Univers, nous permet de contraindre les densités des différentes composantes de l'Univers et notamment les paramètres w_0 et w_a de l'équation d'état de l'énergie noire.

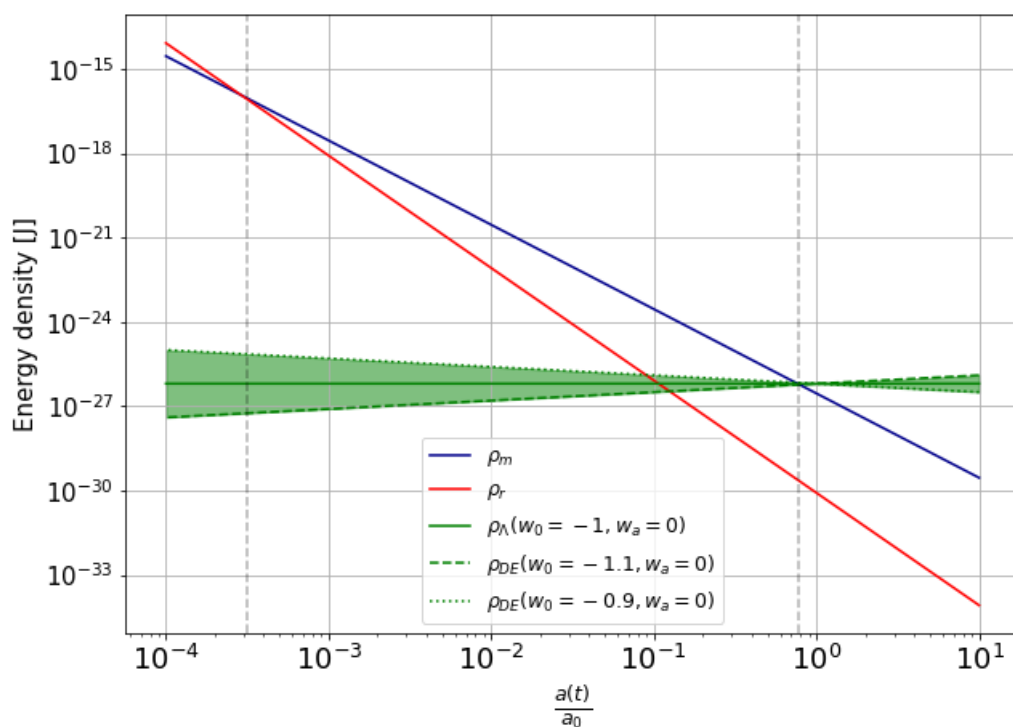


FIGURE 1.12 – Densité d'énergie en fonction du facteur d'échelle pour les différents fluides composant l'Univers, en vert pour l'énergie noire, en bleu pour la matière et en rouge pour la radiation. Trois phases se succèdent, délimitées par les lignes verticales, une première pour laquelle la densité de particules relativistes domine, suivie d'une phase pour laquelle la matière non-relativiste domine. Aujourd'hui la densité d'énergie noire domine.

Quintessence Il existe plusieurs modèles de champs d'énergie noire [Yoo & Watanabe (2012)]. La classe des modèles quintessence est répandue [Caldwell et al. (1998)] et introduit un champ scalaire, noté ϕ , [Tsujikawa (2013)]. Dans la plupart des modèles, ce champ interagit avec les autres seulement par gravitation. Un tel champ variant lentement et suivant un potentiel, noté $V(\phi)$ peut décrire une accélération de l'Univers. Pour cela la masse de quintessence $m_\phi^2 = \frac{d^2V(\phi)}{d\phi^2}$ doit être basse, inférieure en valeur absolue à $H_0\hbar/c^2$.

On peut écrire l'action de ce champ ϕ dans le potentiel $V(\phi)$:

$$S_\phi = \int \sqrt{-g} \left[-\frac{1}{2} g^{\mu\nu} \partial_\mu \partial_\nu \phi - 2V(\phi) \right] d^4x \quad (1.40)$$

où g est le déterminant de la métrique. La pression et la densité d'énergie sont données respectivement, dans la cas de la métrique FRLW pour un Univers plat, par les équations : $P = \frac{\dot{\phi}^2}{2} - V(\phi)$ et $\rho = \frac{\dot{\phi}^2}{2} + V(\phi)$. Ce qui donne pour w :

$$w = \frac{P}{\rho} = \frac{\frac{\dot{\phi}^2}{2} - V(\phi)}{\frac{\dot{\phi}^2}{2} + V(\phi)} \quad (1.41)$$

Dans cette équation, on voit que si le champ varie lentement à une époque tardive, c'est-à-dire $\dot{\phi} \ll V(\phi)$, $w \sim -1$, l'énergie noire se comporte comme une constante cosmologique. L'équation du mouvement pour ce champ est l'équation de Klein-Gordon :

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\phi} = 0 \quad (1.42)$$

Ces modèles prédisent des formules analytiques pour w , facilement testables. Par exemple, les modèles proposés par [TsujiKawa \(2013\)](#) proposent des variations au modèle Lambda-CDM de l'ordre de quelques pourcents sur w_0 . Ces prédictions sont compatibles avec les mesures actuelles, mais donnent une idée de la précision de mesure que devront avoir les prochains relevés. Si la déviation d'une équation d'état constante avec le *redshift* était observée, les modèles de quintessence deviendraient de sérieux candidats pour expliquer l'énergie noire.

1.5.2 Le fond diffus cosmologique

Sur la figure 1.12, on voit à z élevé, que la densité de radiation domine les autres. À cette époque l'Univers est dense et chaud. Photons, électrons et baryons constituent un plasma. Dans ce plasma, le libre parcours moyen des photons est faible, de sorte que la matière est constamment ionisée. La température diminue avec l'expansion de l'Univers, puisque $\rho \propto T^{-4}$. Au fur et à mesure de l'expansion, la densité de rayonnement diminue plus rapidement que la densité de matière, jusqu'à ce que cette dernière domine l'expansion vers $z \sim 3200$.

La température diminue donc jusqu'à ce que les baryons et les électrons se recombinent : c'est la *recombinaison*. Les photons n'interagissent plus ni avec les baryons, ni avec les électrons, maintenant liés en noyaux neutres, l'Univers devient transparent. Jusqu'avant la recombinaison, les interactions entre photons et baryons se font par diffusion Thomson. À la recombinaison ces interactions cessent brutalement laissant une surface de dernière diffusion à $z \sim 1100$. Ces photons sont encore observables aujourd'hui et composent le *fond diffus cosmologique*. Ce rayonnement a été observé pour la première fois et de manière fortuite par Penzias et Wilson en 1965, [[Penzias & Wilson \(1965\)](#)]. En 1994, le satellite COBE a mesuré le spectre du CMB et démontré son incroyable accord avec un spectre de corps noir à $2.7255 \pm 0.0006K$ [[Fixsen \(2009\)](#)] (cf. Figure 1.13), sur plus de quatre ordres de grandeurs en intensité et trois en fréquence. On peut voir sur la figure 1.1, image de la distribution de la matière à l'époque de la recombinaison, l'ordre de grandeur extrêmement faible des anisotropies de température du CMB : $\frac{\Delta T}{T} \sim 10^{-5}$.

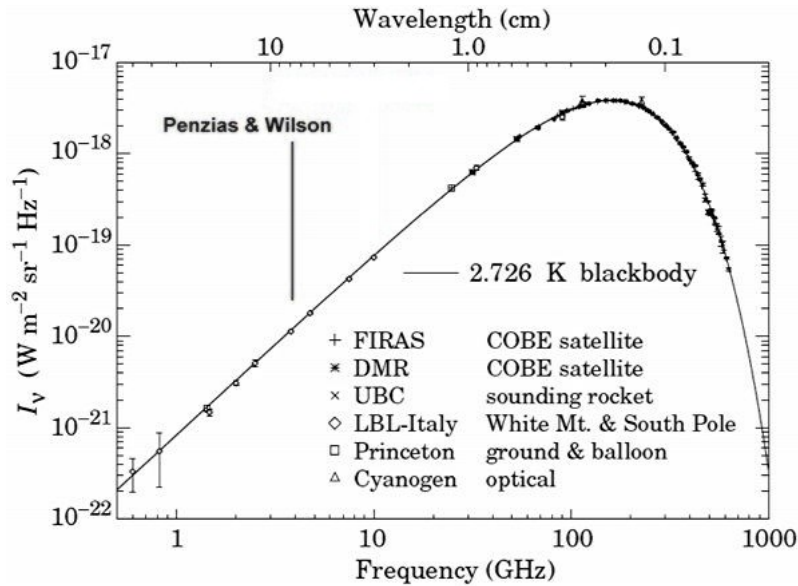


FIGURE 1.13 – Spectre du CMB, les points sont les données mesurées par des expériences spatiales et au sol. La courbe représente le calcul théorique d’un spectre de corps noir à $T \sim 2.7255$ K.

1.5.3 Croissance des structures

Suite à la recombinaison, la matière domine, son effondrement gravitationnel commence et les faibles anisotropies de densité sont amplifiées : par attraction gravitationnelle, les régions denses se densifient et les régions initialement sous-denses le sont de plus en plus. Au fur et à mesure, que la matière s’effondre gravitationnellement, les premières étoiles et galaxies se forment. La galaxie la plus lointaine observée [Oesch et al. (2016)] a été détectée à $z = 11.09$.

À $z \sim 7$ [Planck Collaboration et al. (2020)], se produit la réionisation, c’est-à-dire l’ionisation de l’hydrogène neutre du milieu interstellaire par la grande quantité d’énergie apportée par de nouveaux objets. Les sources ne sont pas connues mais pourraient être les premières étoiles, galaxies et AGN. Le recherche de ces sources et de l’époque exacte de la réionisation sont des domaines actifs de la cosmologie, notamment à travers l’étude de la raie d’hydrogène à 21cm.

L’effondrement gravitationnel a continué pour former l’Univers tel qu’on peut l’observer autour de nous. Cet Univers est composé de grands groupements, ou *clusters* de galaxies, appelés amas. Les plus gros sont appelés super-amas (dont la taille varie entre des dizaines et une centaine de Mpc [Chon et al. (2014)] constituant des filaments de matières formant les grandes et complexes structures présentes sur la figure 1.2. L’étude de la formation de ces structures aux grandes échelles cosmologiques est la encore un domaine actif de recherche en cosmologie. L’étude des grandes structures de l’Univers permet non seulement de contraindre le modèle standard de la cosmologie mais aussi de valider la théorie de la relativité générale aux échelles cosmologiques. Il existe une multitude de différents modèles de gravité modifiée, qui permettent notamment d’expliquer l’accélération de l’expansion de l’Univers sans ajout d’énergie noire ou encore la courbe de rotation des galaxies sans ajout de matière noire ([Joyce et al. (2016)]). En pratique, il existe des sondes liées à la formation des structures qui permettent de contraindre ces différents modèles :

- L’étude de *clusters* [Allen et al. (2011)] permet de reconstituer la densité de matière dans l’Univers et de pouvoir contraindre son évolution. Avec l’observation de milliers voir de millions de galaxies avec les observatoires de nouvelles générations [DESI Collaboration

et al. (2016), LSST DESC (2012), Laureijs et al. (2014)], l'étude des clusters de galaxies sera une sonde importante pour tester la relativité aux échelles cosmologiques.

- Le lentillage gravitationnel [Mandelbaum (2018)] est le phénomène de déviation de la lumière lors de son passage proche d'une source gravitationnelle. On classe les phénomènes de lentillage gravitationnel en deux catégories : les distorsions importantes visibles à l'oeil nu, appelé lentillage fort, et ceux invisibles dont la détection se fait par une étude statistique des galaxies d'arrière plan, autour d'une large masse : le lentillage gravitationnel faible. Cette étude statistique des formes et des alignements des galaxies permet de reconstituer les zones où le potentiel gravitationnel sera plus important et donc de tracer la distribution de matière dans l'Univers.
- La distorsion dans l'espace des *redshifts* est aussi un traceur de la densité de matière dans l'Univers [DESI Collaboration et al. (2016)]. La vitesse particulière des galaxies tombant dans les puits de potentiel des grandes structures, distord la relation distance-*redshift* provoquant des anisotropies de l'espace des *redshifts*, aussi appelé RSD pour *Redshift Space Distorsion*.

Toutes ces sondes traçant la densité de matière dans l'Univers sont autant de tests de la relativité générale, permettant de contraindre le modèle cosmologique.

1.6 L'étude de l'énergie noire avec la combinaison de différentes sondes cosmologiques

La solidité du modèle Λ CDM provient de la cohérence des mesures des paramètres cosmologiques via différentes sondes. L'utilisation de plusieurs sondes différentes permet de combiner des analyses et ainsi de briser les dégénérescences entre les paramètres du modèle cosmologique qu'elles mesurent.

1.6.1 Les chandelles standards : SNe Ia

Si nous utilisons la formule 1.15, pour mesurer la distance de luminosité d_L , il suffit de mesurer le flux d'un objet f . Sur la figure 1.14 se trouve le diagramme de Hubble historique réalisé par Perlmutter et al. (1999). Les observations proviennent de deux expériences. Les points jaunes sont les SNe Ia proches observées lors du suivi Calan Tololo [Hamuy et al. (1993a)] et les rouges, celles du suivi *Supernova Cosmology Project* [Perlmutter et al. (1999)], produisant au total un diagramme de Hubble de 42 SNe distantes, jusqu'à un *redshift* de $z \sim 0.8$. Sur la figure 1.15, sont présentés les contours de confiance dans le plan $(\Omega_\Lambda, \Omega_m)$: la présence d'énergie noire est indispensable pour expliquer les observations, découvrant ainsi l'accélération de l'expansion de l'Univers.

Perlmutter et al. (1999) mesure une valeur pour $\Omega_m \sim 0.28 \pm 0.14$, en comptant les erreurs statistiques et systématiques identifiées et $\Omega_\Lambda \sim 0.7$, démontrant ainsi que l'expansion de l'Univers est accélérée. Le but des expériences suivantes³ a été d'obtenir des mesures plus précises pour les paramètres cosmologiques, en augmentant la statistique et la compréhension des erreurs systématiques. C'est ce qui a été réalisé dans JLA, une combinaison de plusieurs expériences, qui a permis l'extension du diagramme de Hubble jusqu'à un *redshift* de $z \sim 1.2$. Le diagramme de Hubble a ensuite été enrichi dans le cadre de l'analyse Panthéon [Scolnic et al. (2018), Brout et al. (2022)], que l'on détaillera dans les chapitres suivants.

3. Dont l'historique sera détaillé au chapitre 3.

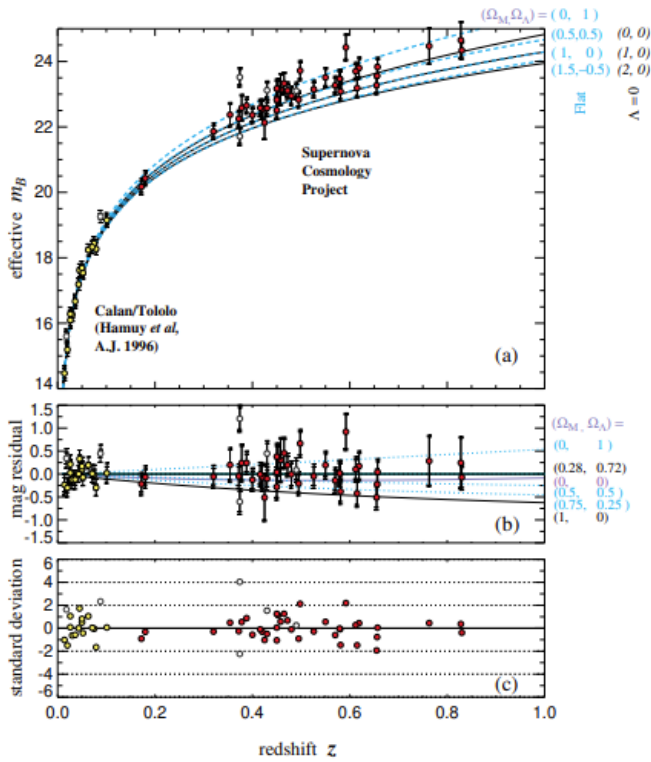


FIGURE 1.14 – Diagramme de Hubble [Perlmutter et al. (1999)] de 42 SNe jusqu'à un *redshift* de 0.8. On voit sur le diagramme de résidus que le modèle privilégié est composé à 72 % d'énergie noire sous forme de constante cosmologique et 28 % de matière.

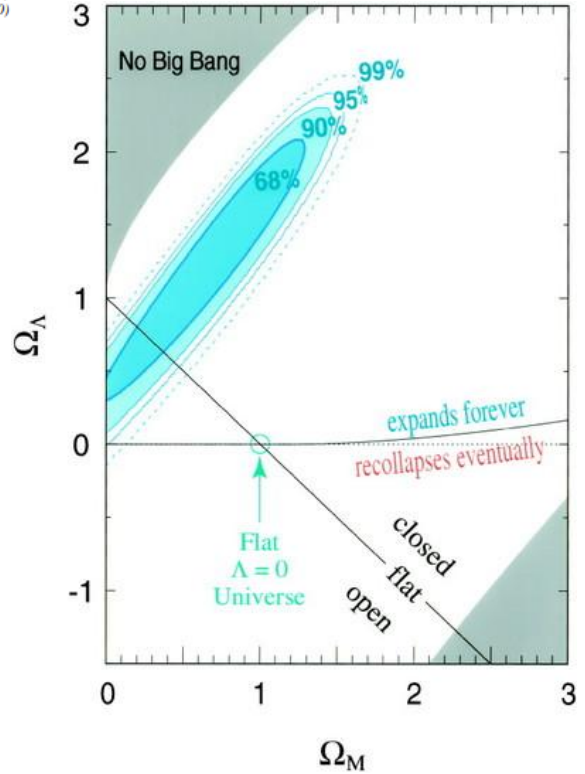


FIGURE 1.15 – Contours de confiance dans le plan $(\Omega_\Lambda, \Omega_m)$, autour de paramètres qui minimisent les résidus de l'ajustement de la figure 1.14.

1.6.2 Le pic de l'oscillation acoustique des baryons

Le fond diffus cosmologique (CMB)

Nous allons nous intéresser aux sources d'anisotropie provenant des inhomogénéités de l'Univers au moment de la recombinaison, en particulier des anisotropies de température. Elles sont générées par les oscillations acoustiques dues aux fluctuations de densité. En effet, lorsque la densité de matière prend le pas sur la densité de radiation, la matière noire commence à s'effondrer, car elle n'est pas couplée au fluide baryons-photons. L'effondrement gravitationnel de la matière noire tente d'entraîner les baryons, liés aux photons. Ils vont produire une pression de rayonnement pour contre-balancer l'effondrement gravitationnel. Ce processus induit une oscillation de densité et la propagation d'ondes acoustiques nommées *oscillations acoustiques baryoniques*, notées BAO pour *Baryonic Acoustic Oscillations*. La phase et la taille des perturbations au moment de la recombinaison sont figées. Ces caractéristiques sont présentes dans les propriétés statistiques de la distribution spatiale des photons du CMB dans ce que l'on nomme les *pics acoustiques*. Les mesures de Planck ont produit les mesures les plus précises du pic BAO et des pics secondaires. Ces propriétés statistiques suivent les symétries de l'Univers. Ainsi, dû à l'isotropie, elles sont indépendantes de la position sur le ciel et invariantes par rotation. Cet aspect associé au fait que

les fluctuations soient gaussiennes, implique que les propriétés statistiques soient complètement décrites par une statistique du second ordre, le *spectre de puissance angulaire*.

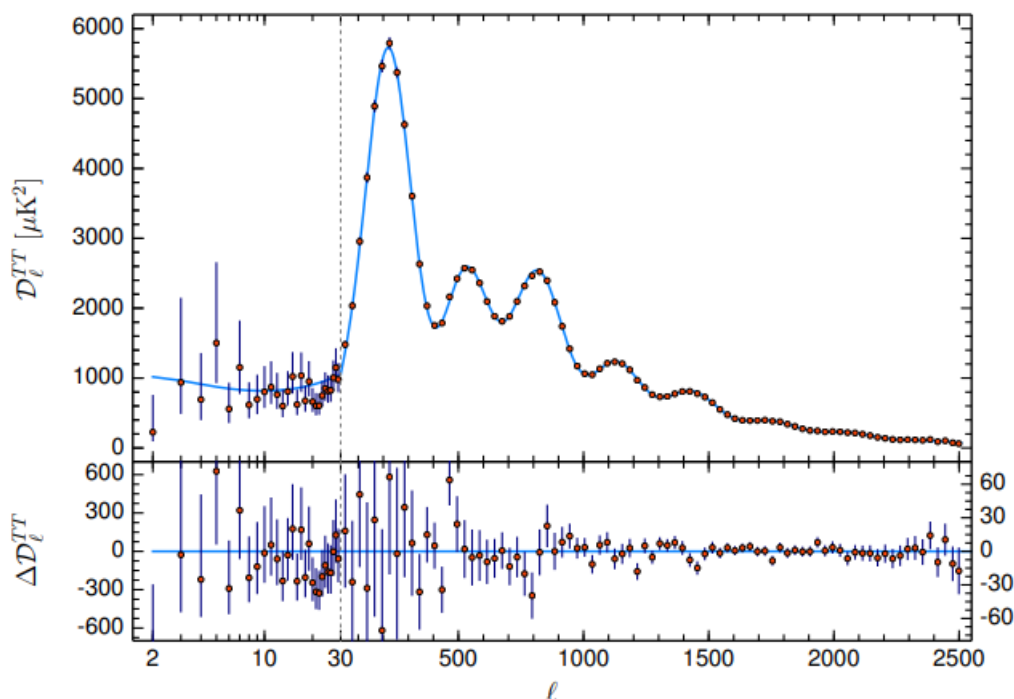


FIGURE 1.16 – Fonction du spectre de puissance des anisotropies de température du CMB (points rouges), modèle Λ CDM ajusté (courbe bleue).

Il existe des fluctuations de température dues principalement à deux effets. D'abord une modification de la température induit par un décalage gravitationnel des photons. Les photons provenant d'une zone de sur-densité sont décalés vers le rouge. Ils doivent traverser une barrière de potentiel, leur faisant perdre de l'énergie. Ceux provenant d'une zone de sous-densité sont décalés vers le bleu. Cet effet est en partie compensé par un décalage temporel dû au fait que les photons provenant de régions sur-denses vont être diffusés plus tôt que ceux provenant de régions sous-denses. Ces deux phénomènes sont regroupés sous le terme d'effets de Sachs-Wolfe. Ensuite un effet Doppler dû à la vitesse propre des électrons au moment de la dernière diffusion, va induire un effet Doppler additionnel. Les photons percutant un gaz s'éloignant d'eux subissent un décalage vers le rouge diminuant la température mesurée dans leur direction.

On peut voir une fonction du spectre de puissance de ces anisotropies sur le graphique supérieur de la figure 1.16 en fonction de l'inverse de l'échelle angulaire, l . On distingue trois régions dans ce graphique. Une première à petit l , à grande échelle, pour laquelle la variance des observations est grande. Nous ne possédons qu'une observation de l'Univers depuis la Terre et la statistique aux grandes échelles est limitée par la variance cosmique (le fait qu'il n'existe qu'un seul Univers). Sur une seconde région, pour $30 < l < 1500$, on distingue une série d'oscillations, les pics acoustiques. Et enfin, pour les petites échelles, les oscillations disparaissent lorsque l augmente puisque le libre parcours moyen des photons augmente.

Observations de la distribution de matière dans l'Univers

Dans le cas du CMB, on mesure donc le pic BAO dans les anisotropies de température, mais on peut aussi retrouver l'empreinte du pic BAO dans la distribution de matière à grande échelle.

Considérons la figure 1.17, sur laquelle on voit l'évolution du profil de masse d'une surdensité composée de matière noire (en noir), de matière baryonique (en bleu), de photons (en rouge) et de neutrinos (en vert), tiré de Eisenstein et al. (2007), illustrant le phénomène de découplage décrit à la section 1.5.2. À l'origine toutes les espèces ont le même profil proche du centre de la surdensité. Le premier graphique en haut à gauche montre, à $z = 6824$, une onde de pression se déplaçant dans plasma baryon-photon couplé et quittant le centre de la perturbation à cause de la pression de rayonnement. L'interaction gravitationnelle liant la matière noire aux baryons, retient le plasma baryons-photons. On peut voir un élargissement des pics. Les neutrino eux se propagent librement. À environ $z = 1100$, a lieu la recombinaison. À $z = 848$, les baryons et les photons sont totalement découplés, impliquant que ces derniers se déplacent librement. Sur la figure en bas à gauche, on peut voir que les photons ont diffusé et les profils de matière noire et de matière baryonique se rejoignent, par interaction gravitationnelle à la position comobile du pic BAO à environ 150 Mpc. Cette position n'a pas évolué depuis la recombinaison. Les mesures de cette position à différentes périodes de l'Univers informent sur l'évolution de taux d'expansion.

En pratique, on veut mesurer l'*échelle acoustique*, la distance comobile entre leur émission et le moment où les photons cessent de ressentir l'effet des baryons. Cet instant est appelé *époque de drag*, qui a lieu peu de temps après la recombinaison, et est noté r_d :

$$r_d = \int_{z_{drag}}^{+\infty} \frac{c_s(z) dz}{H(z)} \quad (1.43)$$

où z_{drag} est le *redshift* de cet instant, $c_s(z)$ est la vitesse du son dans le plasma baryon-photon, qui dépend des densités des deux composantes et donc du *redshift*. r_d est prédit par les mesures du CMB, sa valeur est 147.18 ± 0.29 Mpc ([Planck Collaboration et al. (2020)]). Dans un Univers plat, on peut définir les distances suivantes [Hogg (1999)] :

$$d_M(z) = (1+z)d_a(z) \quad \& \quad d_H(z) = \frac{c}{H(z)} \quad (1.44)$$

L'échelle acoustique peut être mesurée perpendiculairement à la ligne de visée, en mesurant la taille angulaire $\Delta\theta$ de la perturbation. Elle peut aussi être mesurée parallèlement à la ligne de visée, dans ce cas on mesure une différence de *redshift* Δz autour de la perturbation, de *redshift* z :

$$\Delta\theta = \frac{r_d}{d_M(z)} \quad \& \quad \Delta z = \frac{r_d}{d_H(z)} \quad (1.45)$$

En pratique, la mesure des BAO ne contraint pas les paramètres cosmologiques directement. Elle mesure des déviations de certaines distances à un modèle fiduciel, généralement Λ CDM mesuré par Planck, paramétrisées par a_{\parallel} et a_{\perp} :

$$a_{\perp} = \frac{[d_M(z)/r_d]_{mesure}}{[d_M(z)/r_d]_{CMB}} \quad \& \quad a_{\parallel} = \frac{[d_H(z)/r_d]_{mesure}}{[d_H(z)/r_d]_{CMB}} \quad (1.46)$$

On obtient donc deux mesures indépendantes du taux d'expansion de l'Univers au *redshift* z , la première intégrée, la seconde directe. De ces mesures, on peut déduire une valeur de Ω_{Λ} et

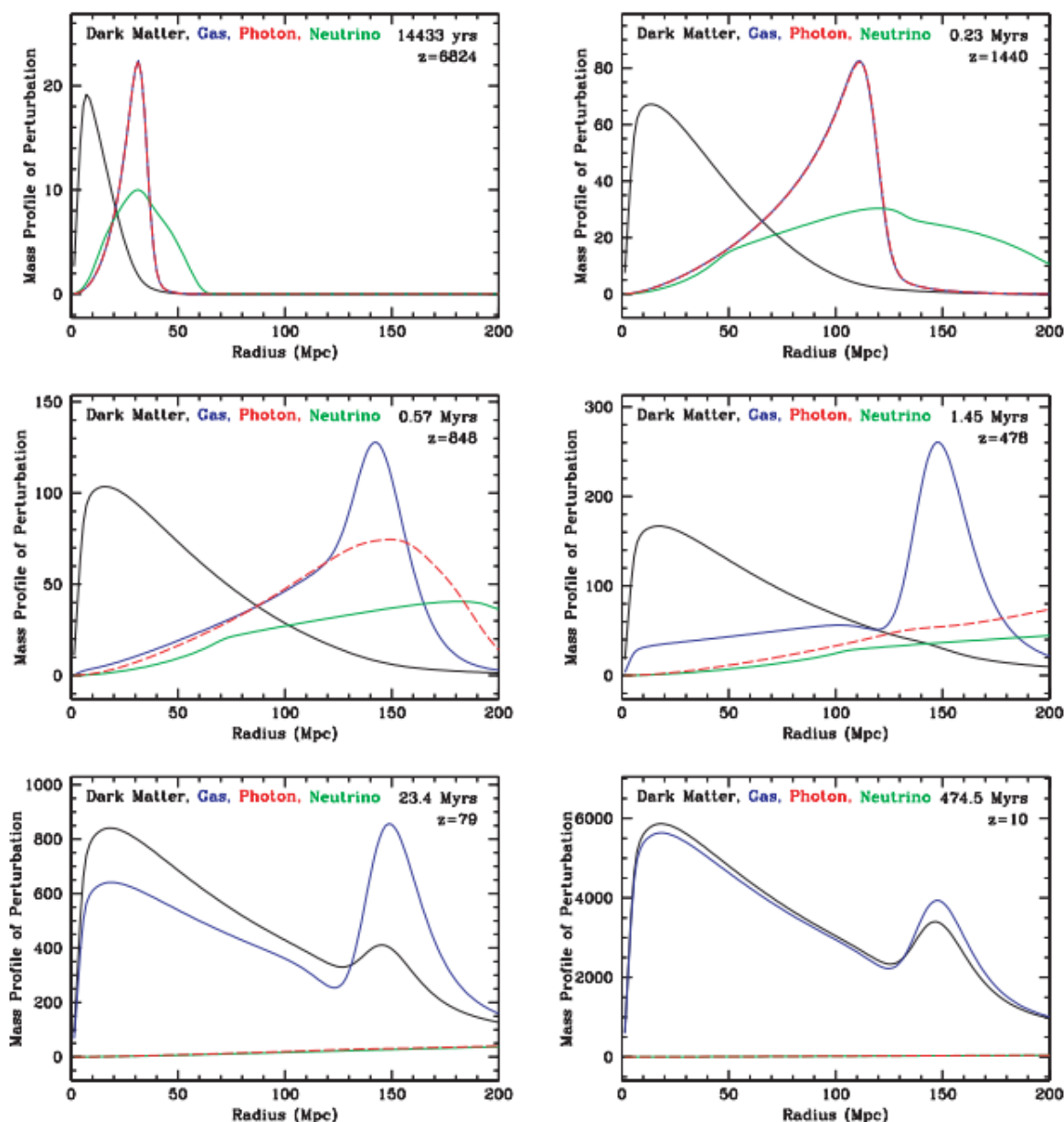


FIGURE 1.17 – Coupe à différents *redshift* de l'évolution du profil de masse de chaque espèce en fonction du rayon comobile, extrait de Eisenstein et al. (2007)

Ω_m . On réalise un diagramme de Hubble (cf. Figure 1.19), permettant de contraindre l'histoire de l'expansion de l'Univers en fonction du *redshift*. Pour mesurer les distances définies par les équations 1.44, on mesure le pic BAO, à différents *redshifts*, impliquant l'observation de différents traceurs. On peut citer trois traceurs de matières :

- les galaxies, première détection du pic BAO, sondent la distribution de matière à bas *redshift*. On remarque que le signal est faible, il faut donc une grande densité d'objet (cf. figure 1.18). On peut tracer la distribution de matière avec différents types de galaxies (les cinq premiers point de la figure 1.19).
- les quasars sont des rayonnements lumineux émis par un trou noir super-massif au centre d'une galaxie [Shields (1999)]. Ils sont extrêmement lumineux, jusqu'à $10^{12}L_\odot$ [Greenstein

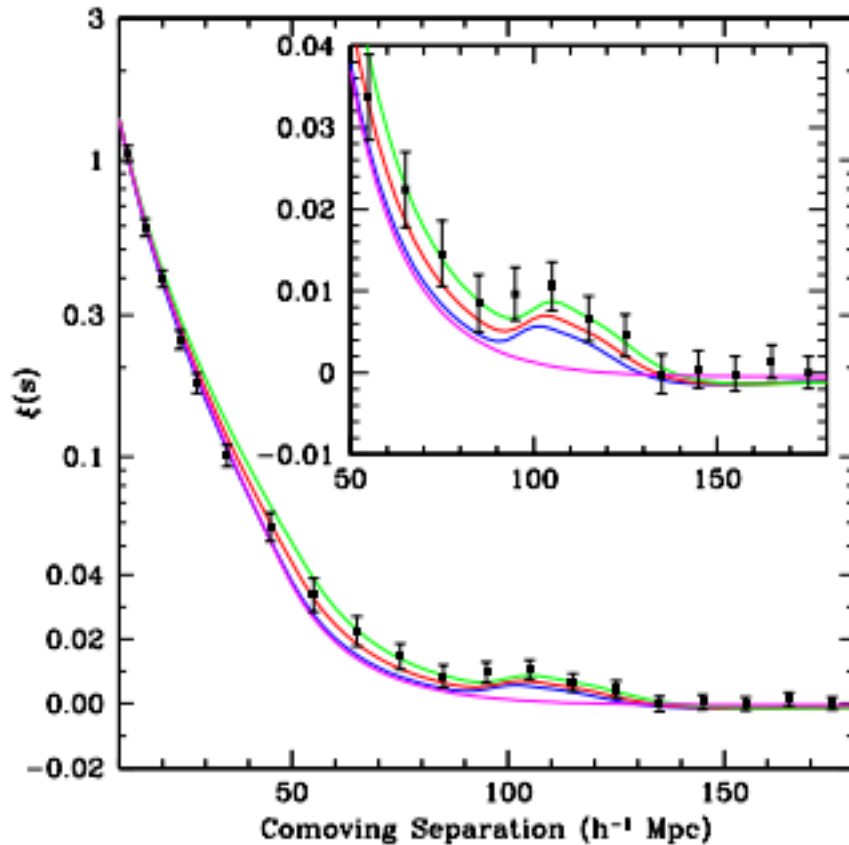


FIGURE 1.18 – Fonction de corrélation dans l'espace des *redshifts* aux grandes échelles du relevé Large Red Galaxies (LRG) du SDSS. Différents modèles cosmologiques possédant différentes valeurs d'équation d'état de l'énergie noire sont représentés en couleur. Le modèle en magenta est un modèle purement CDM, il manque le pic BAO.

- & Schmidt (1964)] et donc visibles à grand *redshift*, $z \sim 5$ (sixième point du diagramme 1.19) ;
- forêt Lyman α , correspondent aux raies d'absorption de l'hydrogène présent dans le spectre des quasars, indicateur de la variation de matière sur la ligne de visée (deux derniers points de la figure 1.19).

1.6.3 Combinaison de sondes

Chacune des différentes sondes cosmologiques permet de caractériser le paramètre d'état, w , de l'énergie noire et leur combinaison permet de s'affranchir des dégénérescences avec les autres paramètres cosmologiques.

Pour briser ces dernières, on combine les différentes méthodes de mesures.

Le CMB et les BAO apportent de fortes contraintes sur la courbure de l'univers, contrairement aux SNe. Ces sondes sont utilisées pour contraindre la courbure de l'Univers à 0. Les modèles d'Univers plats sont ensuite ajustés pour construire un diagramme de Hubble des SNe ayant un fort impact sur l'équation d'état de l'énergie noire. On peut voir sur la figure 1.20, pour un modèle d'Univers plat que le CMB seul (contour vert) n'est pas très sensible au paramètre d'état de l'énergie noire, contrairement aux SNe (contour bleu). Leur combinaison permet d'améliorer

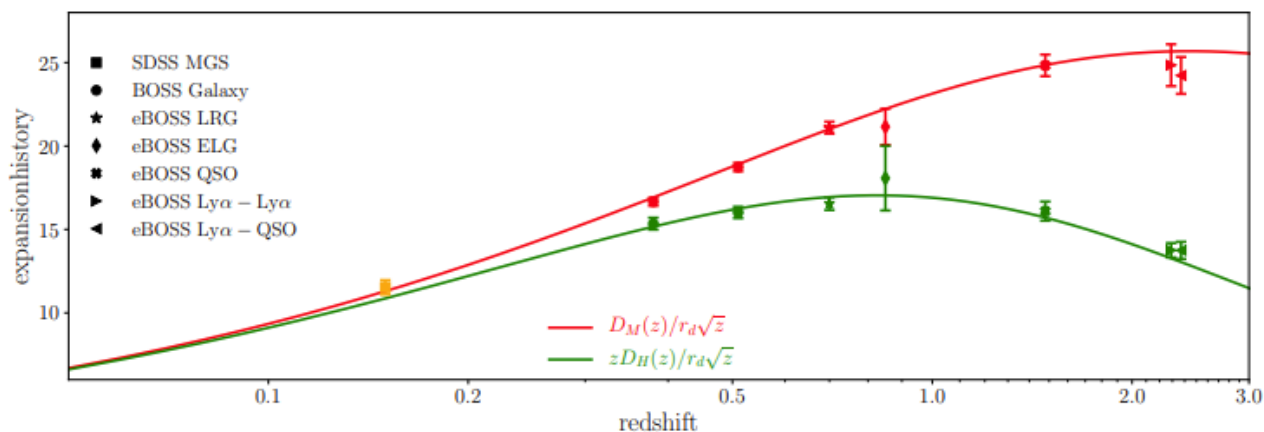


FIGURE 1.19 – Diagramme de Hubble (équation 1.44) réalisé avec différents traceurs de matière observés à différents *redshifts*. Figure extraite de Alam et al. (2021)

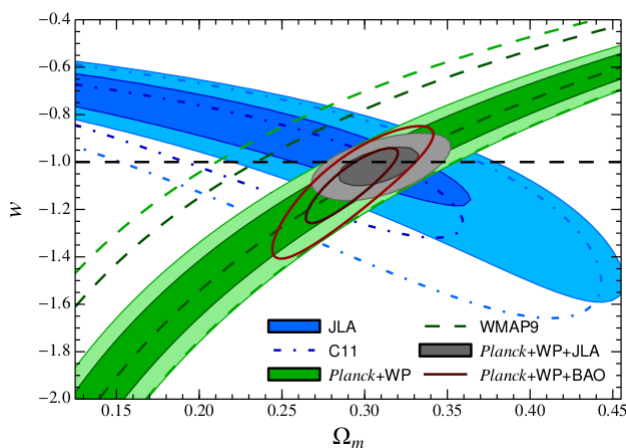


FIGURE 1.20 – Contours de confiance dans le plan (w, Ω_m) extrait de Betoule et al. (2014). La zone en bleu correspond aux contraintes imposées par les SNe, en vert celles produites par le CMB et en gris leur combinaison. Figure extraite de [Betoule et al. (2014)].

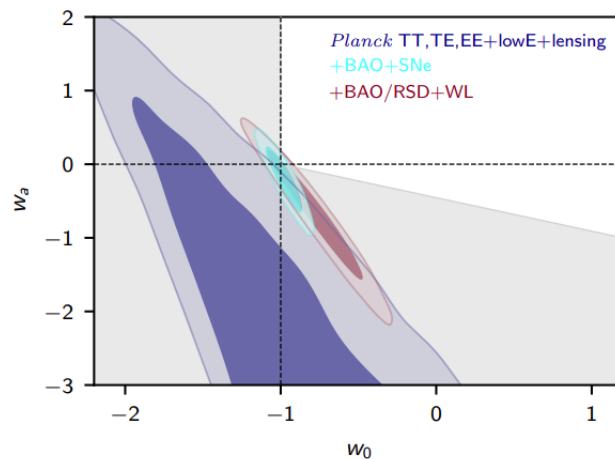


FIGURE 1.21 – Contours de confiance dans le plan (w_a, w_0) extrait de Planck Collaboration et al. (2020), on voit que la contrainte la plus précise sur la mesure de (w_0, w_a) est celle imposée par la combinaison CMB + BAO + SNe

considérablement les contraintes (contour gris). Les courbes rouges correspondent à la combinaison CMB + BAO, qui est plus sensible à la densité de matière dans l'univers et au paramètre d'équation d'état de l'énergie noire. La combinaison la plus contraignante est la combinaison des trois sondes. Sur la figure 1.21, on s'intéresse à la paramétrisation de l'énergie noire décrite équation 1.38. On voit que le CMB seul ne contraint pas les paramètres w_a et w_0 . Ces contraintes sont maximales quand on y ajoute les contraintes imposées par la mesure du BAO et des SNe. Chacune contraint des paramètres différents du modèle cosmologique :

Sonde cosmologique	Observable	Redshift	Paramètres cosmologiques
CMB-T	Anisotropies en température	1100	$A_s, n_s, \Omega_{cdm}/h, \Omega_m/h, \Omega_k/h, H_0$
Distribution de matière	Position des galaxies et quasars	0-5	$\Omega_\Lambda, \Omega_m, \Omega_k, w$
	Absorption spectres de quasars	2-5	
SNe	Flux des SNe	0-2	$\Omega_\Lambda, \Omega_m, w$

Le tableau 1.22 montre les contraintes sur les paramètres cosmologiques imposé par la combinaison des trois sondes différentes de l'énergie noire pour deux cosmologies : $FlatwCDM$ correspond à un Univers plat, avec une équation d'état d'énergie noire de paramètre constant w . En revanche pour le modèle $Flatw_0w_aCDM$, le paramètre d'état de l'énergie noire varie suivant la paramétrisation définie équation 1.38.

La dernière ligne de chaque tableau est ajustement combiné des trois sondes (CMB Planck Collaboration et al. (2020), en supposant un Univers plat, et BAO de BOSS et eBOSS [Ross et al. (2015), Alam et al. (2017), Bautista et al. (2021), Hou et al. (2021), Chabanier et al. (2019)]).

Pour $FlatwCDM$, la combinaison des sondes donne $w = -0.978^{+0.024}_{-0.031}$ et $\Omega_m = 0.316^{+0.022}_{-0.038}$, qui est cohérent avec une énergie noire analogue à une constante cosmologique. Pour $Flatw_0w_aCDM$, les valeurs de $w_0 = -0.841^{+0.066}_{-0.061}$ et $w_a = -0.65^{+0.28}_{-0.32}$ s'écartent à plus de 2σ d'une constante cosmologique. Ce résultat est à mettre en perspectives avec le résultat de Alam et al. (2021), utilisant les mesures finales des BAO et RSD de SDSS et l'analyse Panthéon, et trouvant $w_0 = -0.939 \pm 0.073$ et $w_a = -0.31^{+0.28}_{-0.24}$.

	Ω_M	Ω_Λ	H_0	w_0	w_a
External Probes (No SHOES) - $FlatwCDM$					
Planck & Pantheon+	$0.325^{+0.010}_{-0.008}$	$0.675^{+0.008}_{-0.010}$	$66.49^{+0.50}_{-0.83}$	$-0.982^{+0.022}_{-0.038}$	-
Planck & galaxyBAO & Pantheon+	$0.319^{+0.006}_{-0.007}$	$0.681^{+0.007}_{-0.006}$	$66.78^{+0.76}_{-0.50}$	$-0.974^{+0.025}_{-0.031}$	-
Planck & allBAO & Pantheon+	$0.316^{+0.005}_{-0.008}$	$0.684^{+0.008}_{-0.005}$	$66.87^{+1.00}_{-0.32}$	$-0.978^{+0.024}_{-0.031}$	-
External Probes (No SHOES) - $Flatw_0w_aCDM$					
Planck & Pantheon+	$0.318^{+0.012}_{-0.014}$	$0.682^{+0.014}_{-0.012}$	$67.4^{+1.1}_{-1.2}$	$-0.851^{+0.092}_{-0.099}$	$-0.70^{+0.49}_{-0.51}$
Planck & galaxyBAO & Pantheon+	$0.318^{+0.009}_{-0.006}$	$0.682^{+0.006}_{-0.008}$	$67.12^{+0.71}_{-0.69}$	$-0.878^{+0.063}_{-0.069}$	$-0.45^{+0.29}_{-0.32}$
Planck & allBAO & Pantheon+	$0.316^{+0.009}_{-0.005}$	$0.684^{+0.005}_{-0.009}$	$67.41^{+0.52}_{-0.82}$	$-0.841^{+0.066}_{-0.061}$	$-0.65^{+0.28}_{-0.32}$

FIGURE 1.22 – Résultats cosmologiques réalisé par Brout et al. (2022) de la combinaison des sondes de CMB, BAO et SNe, pour deux cosmologies : $FlatwCDM$ et $Flatw_0w_aCDM$.

1.7 Autres utilisations cosmologiques des SNe Ia

1.7.1 Diagramme de Hubble pour mesurer H_0

La mesure de la constante de Hubble H_0 peut être extraite de l'observation du CMB mais aussi par la mesure de l'échelle absolue du diagramme de Hubble de SNe Ia. L'adéquation de ces mesures, dont la première se base sur la physique de l'Univers primordial et la seconde sur la physique de l'Univers local, permettrait de valider la mesure de l'expansion de l'Univers sur des échelles de temps cosmologique. La contrainte sur H_0 effectuée par Planck est ([Planck Collaboration et al. (2020)]), dans le cadre du modèle ΛCDM :

$$H_0 = 67.4 \pm 0.5 \text{ km.s}^{-1}.\text{Mpc}^{-1}(\text{CMB}) \quad (1.47)$$

Pour mesurer H_0 à partir du diagramme de Hubble des SN Ia, il faut d'abord déterminer leur luminosité absolue. Pour des sources proches, nous pouvons réaliser l'expansion de Taylor de la distance de luminosité et obtenir un polynôme en z , noté $f(z)$. Nous avons $d_L = f(z)/H_0$, ainsi la relation 1.36 devient :

$$m = -2,5 \log(LH_0^2) + 5 \log(cf(z)) + C \quad (1.48)$$

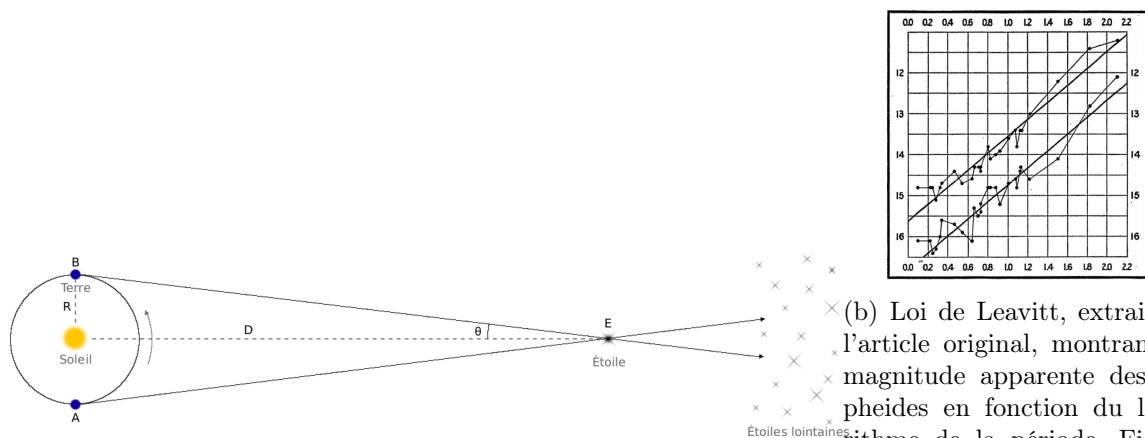
On remarque que la mesure de la magnitude apparente, m , et de *redshift* z ne permet pas d'estimer à la fois la luminosité absolue de la source, et la constante de Hubble mais seulement le produit LH_0^2 . Il est alors nécessaire de briser cette dégénérescence en étalonnant la luminosité absolue de la source utilisée, avec des mesures de distance indépendantes du *redshift*. Pour ce faire, plusieurs techniques peuvent être utilisées.

Utilisation des Céphéides pour calibrer la luminosité des SNe Ia

Il existe peu de façons de mesurer des distances précisément et cela concerne des objets relativement proches. C'est pourquoi on utilise des mesures de proche en proche, technique appelée "échelle de distance" (*distance ladder*), chaque méthode de mesure correspondant à un "barreau" de cette échelle. Cette technique est résumée sur la figure 1.23, présentant l'utilisation des Céphéides pour déterminer la luminosité des supernovæ de type Ia.

Dans un premier temps, à l'échelle de la Galaxie, il est possible de mesurer des distances de manière géométrique par parallaxe, (cf. Figure 1.24a), lors de la rotation de la Terre autour du Soleil, passant du point A au point B, la position de l'étoile E observée depuis la Terre se déplace sur l'arrière plan des étoiles lointaines fixes. L'angle de déplacement θ dépend directement de la distance D nous séparant de l'étoile et du rayon de la Terre R : $D = R/\theta$.

Les Céphéides sont des étoiles variables et la variation temporelle périodique de leur luminosité, T_{Ceph} , est reliée à leur luminosité intrinsèque L_{Ceph} par la loi de Leavitt (cf. Figure 1.24b). La mesure de leur période permet donc d'en déduire leur distance, à condition d'avoir calibré cette relation sur des céphéides suffisamment proches et situées dans la Voie Lactée pour disposer d'une mesure de leur distance par parallaxe. Ces mesures de distance de céphéides dans la Voie Lactée correspondent au premier ensemble de points noté (*Milky Way*) sur le premier cadran de la figure 1.23.



(a) Méthode de mesure de distance en utilisant les parallaxes. Figure extraite de [par (2004)]. (b) Loi de Leavitt, extrait de l'article original, montrant la magnitude apparente des céphéides en fonction du logarithme de la période. Figure extraite de [Leavitt & Pickering (1912)].

FIGURE 1.24 – Illustration de la mesure de distance par parallaxe et loi de Leavitt.

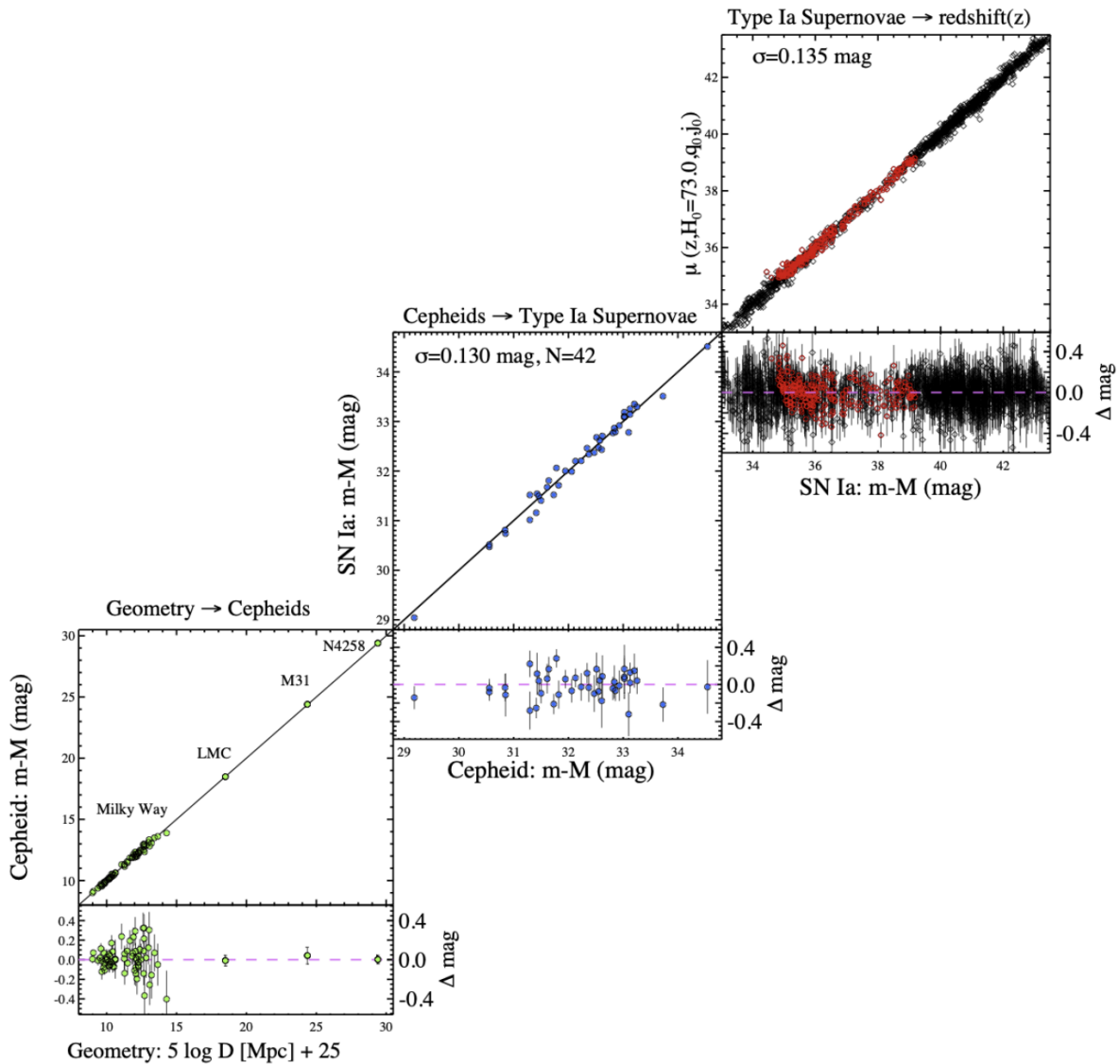


FIGURE 1.23 – Diagramme de Hubble dont le point zéro est déterminé par une échelle de distance. La distance des Supernovæ est calibrée sur les céphéides. Figure extraite de [Riess et al. (2021)].

Pour mesurer la distance des Céphéides extra-galactiques (dans le LMC pour *Large Magellanic Cloud* mais aussi dans M31, figure 1.23), on peut utiliser les systèmes de deux étoiles binaires détachées à éclipses. Le système est en rotation : à partir de la mesure de la vitesse de rotation on détermine le diamètre de l'orbite, qui, comparé à leur angle de séparation, permet d'en déduire leur distance. Cette distance peut être reliée directement à la distance entre le système et nous. Cette observable nous permet de mesurer des distances plus lointaines que les parallaxes. Enfin à des distances plus importantes encore, une troisième technique peut être exploitée : l'observation de megamaser, intense rayonnement radio, dû à la rotation d'un disque autour du trou noir central d'une galaxie. En effet l'excitation de molécules d'eau du disque peut engendrer des megamasers de forte luminosité. Le décalage doppler des raies observées permet d'étudier la dynamique du disque et d'en déduire sa taille, pouvant être comparée à son diamètre apparent.

Le deuxième quadrant utilise des supernovæ de type Ia, dont la luminosité L est quasi identique d'un individu à l'autre. Afin de calibrer cette luminosité, on s'appuie sur l'observation au sein

d'une même galaxie de SNe Ia et de Céphéides. La mesure de distance est réalisée en mesurant la période des Céphéides, et en utilisant la loi de Leavitt calibrée dans le premier quadrant.

Le troisième cadran utilise des supernovæ plus lointaines, de *redshift* $z > 0.01$, mais suffisamment pour réaliser l'expansion de Taylor. La luminosité L des SNe Ia étant fixée par le deuxième quadrant, on peut déduire de ces mesures la valeur de H_0 ([Riess et al. (2021)]) :

$$H_0 = 73.04 \pm 1.04 \text{ (stat + sys) km.s}^{-1}.\text{Mpc}^{-1} \quad (\text{Céphéides + SNeIa}) \quad (1.49)$$

Tip of the Red Giant Branch (TRGB)

Il existe une autre manière d'obtenir la luminosité intrinsèque des SNe, non plus en utilisant les céphéides mais les géantes rouges pour calibrer la luminosité des SNe. En effet, il existe une limite physique supérieure à la taille d'une géante rouge. La physique de ces objets étant connue, on sait d'après le diagramme de température-luminosité, dit diagramme de Hertzsprung-Russell (HR), que la branche de géante rouge atteint un sommet, au delà duquel ces objets ne sont plus stables et ne peuvent donc pas exister. Or cette limite est la même dans toutes les galaxies. On peut donc calibrer la luminosité maximale de l'extrémité de la bande des géantes rouges (TRGB) dans des galaxies de distance connue, puis en observant une galaxies contenant des géantes rouges au sommet et des SNe, on peut déterminer la luminosité intrinsèque des SNe. Cette méthode [Freedman et al. (2019)] donne une valeur pour la constante de Hubble :

$$H_0 = 69.8 \pm 0.8(\pm 1.1\% \text{ stat}) \pm 1.7(\pm 2.4\% \text{ sys}) \text{ km.s}^{-1}.\text{Mpc}^{-1}(\text{TRGB + SNeIa}) \quad (1.50)$$

Distance au CMB (Inversed distance ladder)

Une autre façon d'obtenir la luminosité intrinsèque des supernovæ est de mesurer leur distance, non plus avec une succession de mesure de distance entre l'Univers proche et les SNe, mais une échelle inversée pour laquelle on part de l'Univers primordial et on remonte jusqu'aux SNe. On utilise la mesure du CMB pour prédire l'échelle de l'horizon sonique à la recombinaison, déterminant la mesure de l'échelle BAO. On utilise ensuite cette connaissance dans une région où l'on observe des SNe et des mesures de l'échelle BAO pour calibrer le diagramme de Hubble des SNe. Cette mesure tiré de Macaulay et al. (2019) vaut :

$$H_0 = 67.8 \pm 1.3 \text{ km.s}^{-1}.\text{Mpc}^{-1} \quad (\text{BAO + SNeIa}) \quad (1.51)$$

1.7.2 Vitesses particulières

La vitesse particulière des galaxies est la vitesse propre de ces objets qui n'est pas due à l'expansion de l'Univers. Elle est responsable des distorsions observées dans l'espace des *redshifts*. La dispersion de ces vitesses particulières et des RSD qui en résulte dépendent de la distribution de

matière locale. En utilisant les résidus au diagramme de Hubble des supernovæ de type Ia, on peut inférer les vitesses particulières des galaxies sans étudier les RSD. La dispersion des vitesses particulières des SNe Ia est de 150 km.s^{-1} . La vitesse particulière d'une supernova dépend de la distribution de matière au travers de l'équation de continuité [Graziani et al. (2020)] :

$$\Delta v(x, a) = -aH(a)f(a)\delta(a) \quad (1.52)$$

où Δv est la vitesse propre de la galaxie, x est sa position, a est le facteur d'échelle, H est la loi de Hubble, δ est le champ de matière et f est le facteur d'échelle linéaire. Il décrit l'évolution linéaire des perturbations de densité. En pratique, les supernovæ utilisées doivent être assez proches pour avoir une contribution non négligeable de leur vitesse particulière à leur *redshift*. Le *redshift* observé, z , peut être relié au *redshift* cosmologique dû à l'expansion de l'Univers z_{cosmo} , par la vitesse particulière de la galaxie v par la formule suivante :

$$1 + z = (1 + z_{\text{cosmo}}) \left(1 + \frac{v}{c}\right) \quad (1.53)$$

Le *redshift* observé est déterminé à partir de la galaxie hôte, le *redshift* cosmologique est inféré à partir du diagramme de Hubble et la vitesse particulière permet de contraindre directement les paramètres de formation des structures, d'après l'équation 1.52. Cette sonde peut être particulièrement efficace pour contraindre les modèles de gravité modifiées à bas z , où la variance cosmique est élevée, rendant les autres sondes (RSD) peu contraignantes. Elle nécessite néanmoins une grande statistique, non disponible aujourd'hui. Les futures expériences de détection et de suivi de SNe pourront la rendre accessible.

1.8 Succès & tensions du modèle

En somme, Λ CDM est le modèle le plus simple en accord avec les observations provenant de différentes sondes, permettant une description simple de l'Univers et de son évolution. Il reste cependant des tensions entre l'estimation de *certain*s paramètres par les différentes sondes. La tension sur le paramètre H_0 est la plus critique. La mesure directe, présentée section 1.7.1, mesure grâce aux céphéides et aux SN Ia une déviation de plus de 5σ entre la valeur de $H_0 = 73.04 \pm 1.04 \text{ km.s}^{-1}.\text{Mpc}^{-1}$ [Riess et al. (2021)] et la valeur estimée par Planck [Planck Collaboration et al. (2020)] se basant sur la physique de l'Univers primordial, $H_0 = 67.4 \pm 0.5 \text{ km.s}^{-1}.\text{Mpc}^{-1}$. L'utilisation des TRGB comme calibrateur de la luminosité des SNe montre une diminution de H_0 vers une valeur plus proche de celle mesurée par le CMB, $H_0 = 69.8 \pm 1.9 \text{ km.s}^{-1}.\text{Mpc}^{-1}$. Cette mesure montre un écart de 1σ avec la mesure du CMB. Enfin la mesure de H_0 en ajustant le diagramme de Hubble sur l'information mesurée par la CMB est plus faible que la mesure faite à partir de l'Univers local (cf. équation 1.51). Cette divergence peut s'expliquer par des erreurs systématiques mal comprises dans les différentes analyses ou indique la piste d'une nouvelle physique. Les mesures utilisant les supernovæ de type Ia vues à la section précédente doivent encore être précisées et mieux comprises pour s'accorder vers une même estimation de ce paramètre. Néanmoins, cette tension est corroborée par d'autres sondes (cf. Figure 1.25) qui elles aussi mesurent des valeurs de H_0 supérieures à celle de Planck. Un intermédiaire pourrait être fourni par les ondes gravitationnelles mais la statistique est pour l'instant insuffisante. La porte est ouverte à de la nouvelle physique dont une revue sur les solutions possibles à cette tension peut être trouvée dans Di Valentino et al. (2021). Une seconde tension est celle sur le paramètre $S_8 = \sigma_8 \left(\frac{\Omega_m}{0.3}\right)^{1/2}$, quantifiant l'amplitude des fluctuations de matière dans l'Univers local. Sa

valeur inférée par le lentillage faible, le *clustering* des galaxies et le RSD est plus faible de 2 à 3 σ que la valeur inférée par Planck [Planck Collaboration et al. (2020)] de 0.834 ± 0.016 .

Malgré ces tensions le modèle privilégié par les observations des différentes sondes cosmologiques et par leur combinaison est le modèle Λ CDM comportant six paramètres. Nous sommes entrés dans une ère de cosmologie de précision. Avec les futurs grands relevés cosmologiques de 4^{ème} génération, DESI [DESI Collaboration et al. (2016)], LSST [LSST DESC (2012)], Euclid [Laureijs et al. (2014)], Roman [Spergel et al. (2015)], les mesures des paramètres cosmologiques vont devenir plus précises, grâce à une statistique nettement supérieure. C'est pourquoi il faut comprendre les différentes sources d'erreur systématiques affectant les différentes analyses, pour pouvoir conclure si le modèle Λ CDM est toujours privilégié par les observations ou si les tensions que nous observons entre les différentes sondes sont les traces d'une nouvelle physique. Dans la suite de ce document nous présentons une étude portant sur la compréhension des erreurs systématiques de l'analyse cosmologique réalisée avec les SNe Ia. Dans la suite nous allons détailler la physique de ces objets et comment leur utilisation permet de contraindre les paramètres cosmologiques.

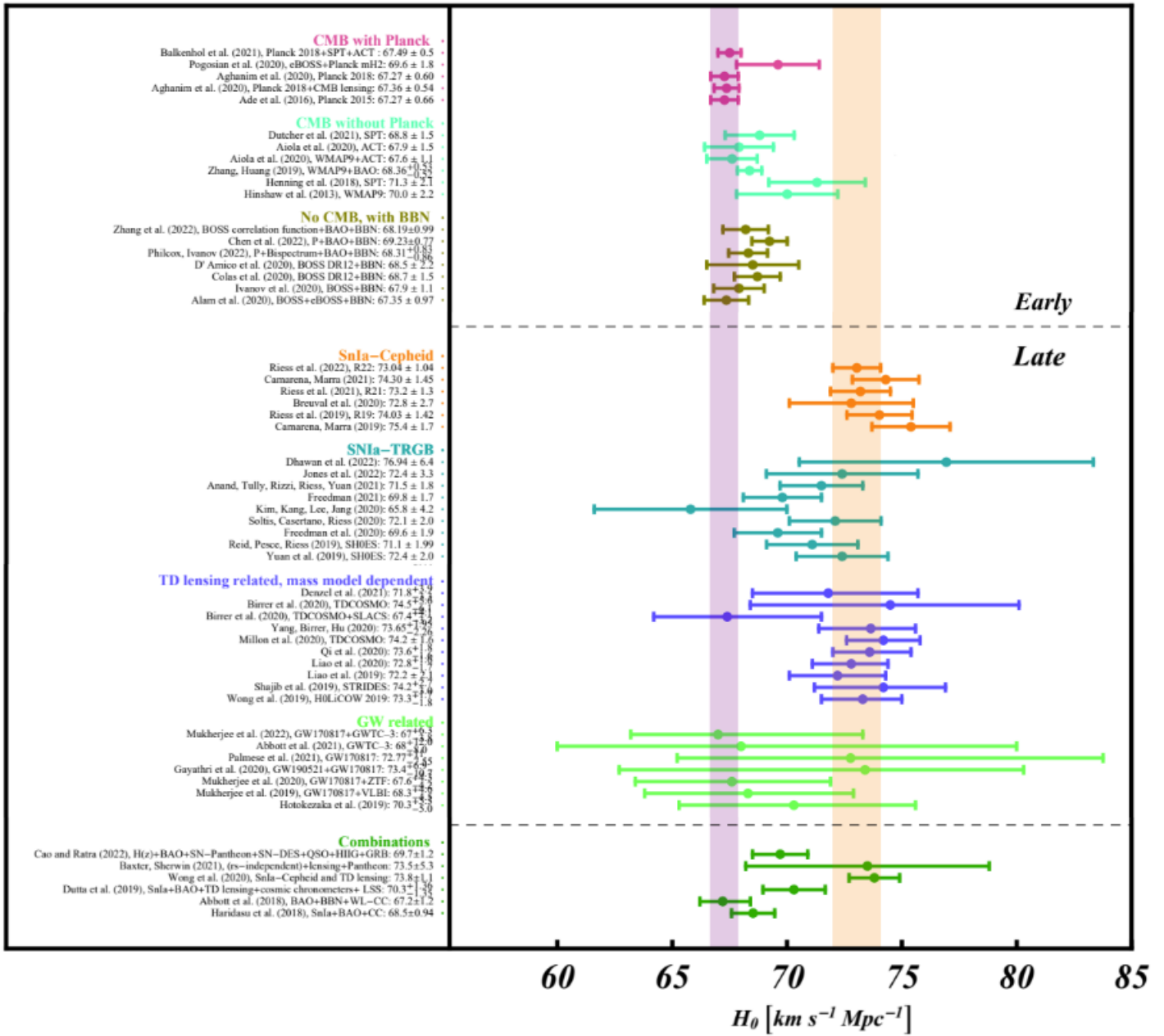


FIGURE 1.25 – Tension sur la valeur de H_0 mesurée par différentes sondes cosmologiques. Les colonnes violette et orange, correspondent respectivement aux valeurs mesurées par Planck et le programme SHOES [Riess et al. (2021)]. Les mesures de l’Univers primordial, se concentre autour de la valeur de Planck dans la colonne violette. Celles inférées par des sondes de l’Univers proche, autour de $74 km.s^{-1}.Mpc^{-1}$, se concentrent dans la colonne orange, on peut notamment citer les SNe Ia mais aussi les mesures de lentillage gravitationnel fort. Figure adaptée de Perivolaropoulos & Skara (2021).

Les Supernovæ de type Ia

Il est possible de contraindre l'équation d'état de l'énergie noire à travers son effet sur l'expansion de l'Univers en mesurant la distance et la vitesse apparente d'éloignement d'objets à différentes périodes de l'Univers. Pour cela, nous devons disposer d'objets ayant une luminosité intrinsèque similaire entre les individus et observables à de grandes distances, c'est à dire à haut *redshift*. C'est la cas des supernovæ de type Ia (SNe Ia). Ces objets extrêmement lumineux, d'environ $10^{10}L_{\odot}$, nous permettent de mesurer des distances cosmologiques. Les SNe Ia sont des objets complexes de luminosité non standard mais standardisable. Certaines caractéristiques observationnelles des SNe Ia doivent être prise en compte pour décrire leur diversité.

Dans ce chapitre, nous verrons comment sont définies les supernovæ et quelles sont les mécanismes compris à l'origine de ces événements. En particulier, nous nous intéresserons à la physique, aux propriétés de supernovæ de type Ia et aux lois qui permettent de diminuer la dispersion de leur luminosité.

Contents

2.1	Observations et mesures des supernovæ	51
2.2	Les Supernovæ (SNe)	51
2.2.1	Classification	52
2.2.2	Proportions & taux d'explosion	54
2.2.3	Explosion par effondrement gravitationnel	55
2.2.4	Supernova thermonucléaire	56
2.3	Propriétés observationnelles des SNe Ia	58
2.3.1	Propriétés spectrales	58
2.3.2	Propriétés photométriques	60
2.3.3	Standardisation	61
2.3.4	Propriétés de la galaxie hôte	64
2.3.5	Sous-classes	66
2.4	Conclusion	67



(a) Transcription d'une observation de nouvelle étoile dans la constellation de Cassiopée, par Tycho Brahe en 1572.

(b) SN 1994D observée par le télescope Hubble [crédit NASA], la luminosité de la SN est similaire à celle de sa galaxie hôte.

FIGURE 2.1 – Deux observations de Supernovæ à quatre cent ans d'intervalle.

2.1 Observations et mesures des supernovæ

Les supernovæ sont des phénomènes stellaires observés en lumière visible. Il existe différents types de supernovæ (SNe). Leur caractérisation s'est faite de manière empirique. Les deux méthodes d'observation usuelles sont la spectroscopie et la photométrie, définies respectivement comme suit.

Spectroscopie La spectroscopie regroupe l'ensemble des procédés permettant d'obtenir le flux d'un objet en fonction de la longueur d'onde à la manière d'un prisme (cf. Figure 2.2). La *résolution* spectrale d'un spectre correspond à la capacité de résoudre de petites structures. Pour chaque observation, un spectre décrivant l'évolution du flux en fonction de la longueur d'onde est obtenu. Plus la résolution spectrale voulue est importante, plus la quantité d'informations à obtenir est élevée et nécessite un temps d'acquisition important. Il limite le suivi spectral des supernovæ les plus éloignées.

Photométrie large bande La photométrie consiste à faire l'image d'une partie du ciel et à mesurer le flux des objets qui s'y trouvent. Ces observations se font dans le domaine visible mais il est utile de le diviser en sous intervalles de longueurs d'onde. En pratique on utilise des filtres afin de pouvoir, par comparaison d'observations dans différents filtres, aussi appelés bandes d'observation, obtenir une information spectrale à basse résolution. Puisque les supernovæ sont des phénomènes variables, elles sont suivies avec une cadence de quelques jours par un instrument pendant quelques mois. Ce suivi permet de constituer l'évolution temporelle du flux observé dans une bande, appelée *courbe de lumière* (cf. Figure 2.3).

2.2 Les Supernovæ (SNe)

2.2.1 Classification

Les astronomes de différentes civilisations ont observé l'apparition soudaine de nouvelles étoiles. En 1572, Tycho Brahe nomme ce phénomène, *nova stella*, du latin nouvelle étoile (cf. Figure 2.1a). Nom qui sera ensuite repris par Johannes Kepler au début du XVII^e siècle. Les travaux de Tycho Brahe ont été d'une grande importance dans l'histoire des sciences, car ils ont notamment permis la déconstruction de la représentation classique du ciel et nourri la révolution Copernicienne qui suivra. Le terme de *Supernova* à lui été introduit en 1931, pour classer les plus lumineuses novæ observées. Cette période, grâce aux avancées de physique fondamentale de l'entre-deux-guerres, a été fructueuse pour la compréhension des objets astrophysiques. On peut citer entre autre la compréhension des mécanismes expliquant la formation des métaux¹ au sein des étoiles, ou encore, la compréhension de la physique des naines blanches grâce à la physique quantique ou la prédiction de trous noirs grâce à la relativité générale. Le terme "supernova" regroupe l'ensemble des phénomènes qui décrivent l'explosion d'étoiles en fin de vie. Ce sont des processus avec une durée de vie finie, observable à l'échelle d'une année, faisant partie de la famille des objets transitoires. Leur flux est initialement faible et augmente pour atteindre un maximum puis diminue jusqu'à devenir indétectable depuis la Terre. La classification des supernovæ s'est faite de manière empirique, dirigée par les observations. Les mécanismes théoriques sont, souvent, encore incertains tant la physique des phénomènes responsables de ces événements est complexe. Les supernovæ ont donc été regroupées, selon leur caractéristiques spectrales dans le tableau suivant définissant le type de la supernova (I ou II) et leur sous-type (par une lettre). Les deux grands types sont les types I et II, ils ont été introduits en 1941 par Minkowski à partir de 14 observations. Le plus gros groupe était composé de 9 objets donc a été nommé I. Une supernova étant une explosion d'étoile, de la matière en est éjectée. Les raies de diffusion que nous voyons dans le spectre sont dues à la diffusion de la lumière provenant de l'étoile par cet éjecta. On peut ainsi voir quels éléments chimiques le composent. Les objets du premier groupe ne présentent pas de raie d'Hydrogène dans leurs spectres contrairement à ceux du second groupe. Sur l'arbre, on peut voir les sous classifications. Le type I se divise en trois sous-types, a, b et c. Ces distinctions ont été établies en 1985 [Wheeler & Levreault (1985) & Elias et al. (1985)] pour les sous-classes a et b, puis en 1990 [Wheeler & Harkness (1990)] pour le sous-groupe c. Puisque la distinction entre ces objets s'est faite observationnellement, avant que la physique sous-jacente ne soit comprise, d'autres classifications sont possibles, notamment par le mécanisme d'explosion qui les engendre. Le premier, commun aux SN Ib, Ic et II, est une effondrement gravitationnelle, en bleu sur le graphique. Le second ne concerne que les supernovæ de type Ia produit d'explosions thermonucléaires, représenté en rouge.

1. En astrophysique, le terme métaux renvoie à tous les éléments plus lourds que l'Hélium.

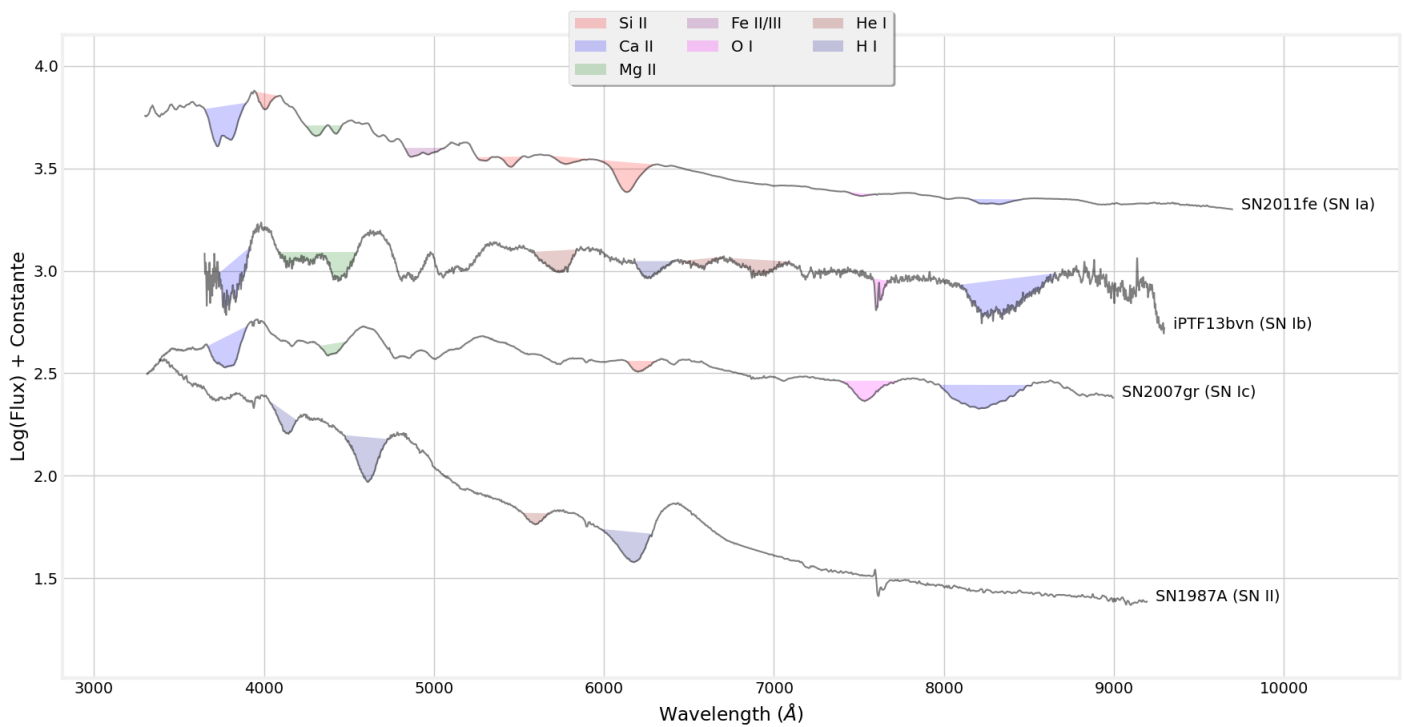
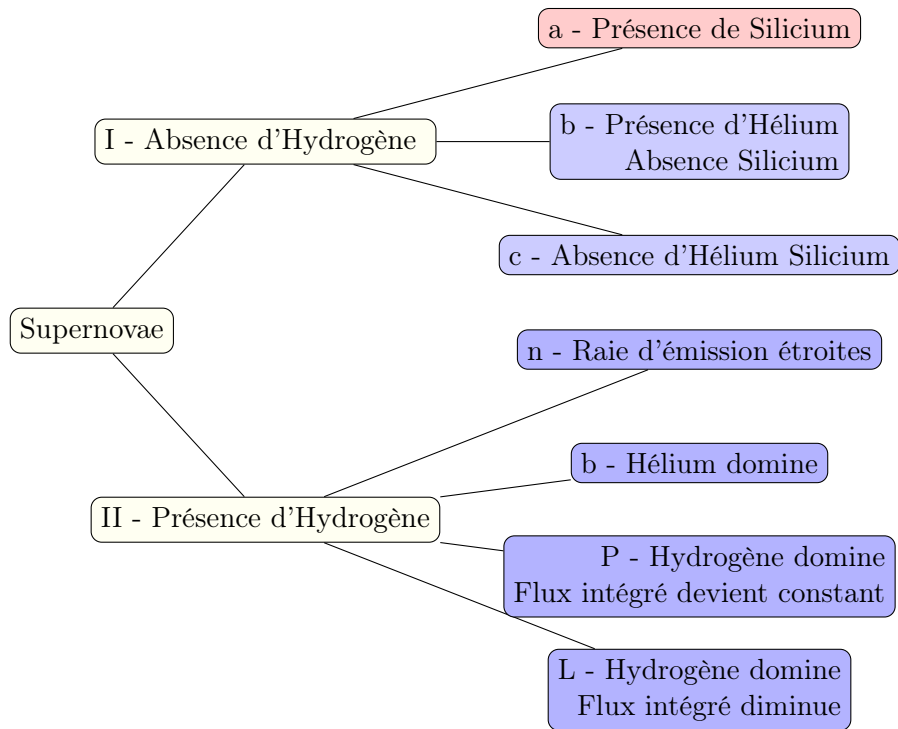


FIGURE 2.2 – Spectres des quatre différents types de supernovæ (de haut en bas) SN Ia, SN Ib, SN Ic et SN II, avec les raies des différents éléments chimiques caractéristiques.

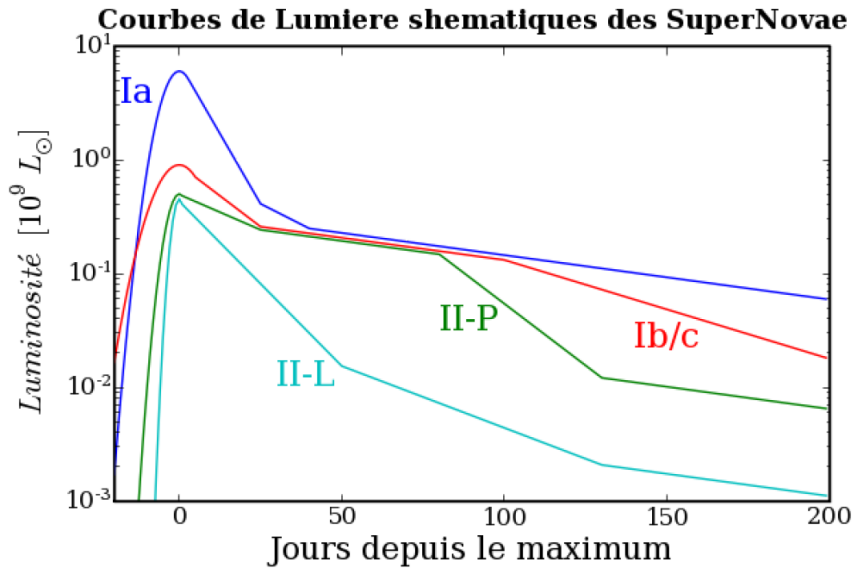


FIGURE 2.3 – Courbes de lumière typiques des grands types de supernovæ. En bleu, la courbe de lumière d’une SN Ia, la luminosité maximale est jusqu’à 10 fois supérieure à celle des autres types de SN. Figure extraite de [Baumont \(2010\)](#).

2.2.2 Proportions & taux d’explosion

Les supernovæ Ia sont les plus lumineuses et de ce fait constituent 79% des SNe observées. Cependant, d’après [Li et al. \(2011\)](#), ce ne sont pas les objets majoritaires dans un volume fini, proche de la Terre. Comme on peut le voir sur le graphique suivant, tiré de [Li et al. \(2011\)](#), 57% des SNe sont des SN de type II dont la luminosité est plus faible, puis 24% de SN Ia et enfin 10% des type Ib ou Ic. Le taux d’explosion des SNe Ia est une donnée importante liée au taux de formation stellaire des galaxies ou encore à la composition du milieu interstellaire. Sa connaissance exacte est indispensable pour contraindre les modèles de progéniteurs, mais sa mesure est compliquée par le fait que ce sont des événements rares.

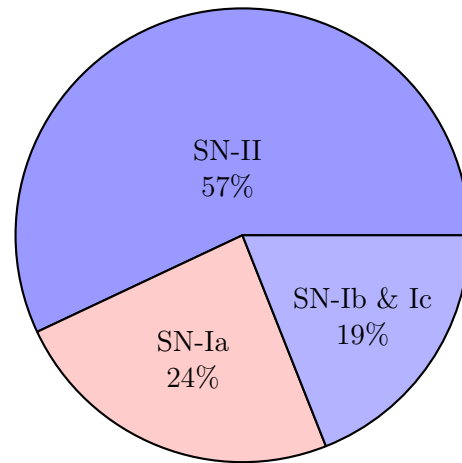


FIGURE 2.4 – Proportions de SNe dans un volume fini [[Li et al. \(2011\)](#)].

Ce taux varie en fonction du *redshift*, puisque la composition de l’Univers et des galaxies évolue elle aussi avec le *redshift*. Sur la figure 2.5, on peut voir le taux d’explosion empirique mesuré par [Frohmaier et al. \(2019\)](#), dont la formule est la suivant :

$$SNRate(z) = (2.27 \pm 0.19)(1+z)^{1.70 \pm 0.21} \times 10^{-5} \text{ SNe}_{Ia} \text{ an}^{-1} \text{ Mpc}^{-3} h_{70}^{-3} \quad (2.1)$$

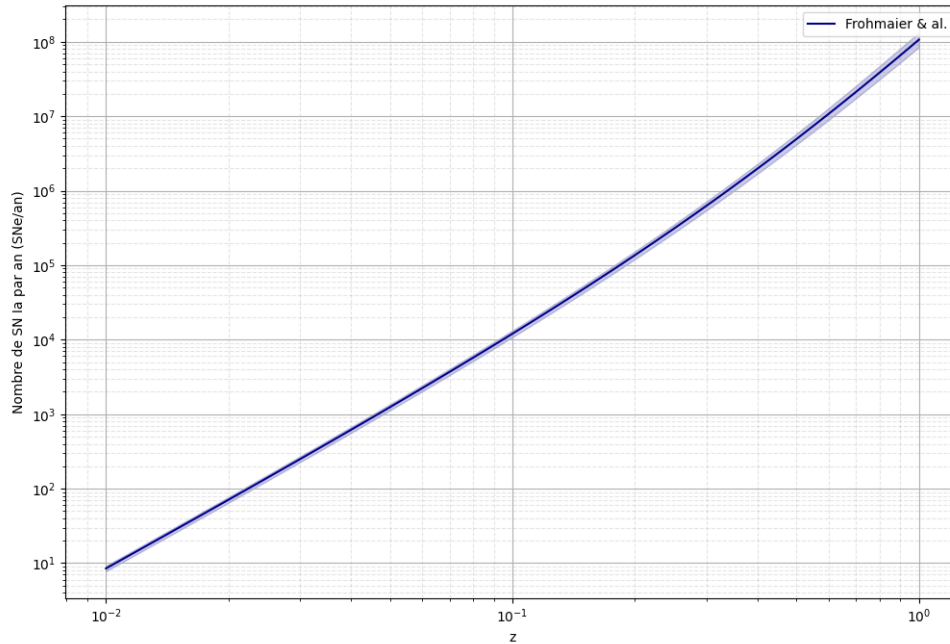


FIGURE 2.5 – Taux d’explosion de SN Ia obtenu en combinant les données de différents relevés de SNe : SDSS, SNLS, SUBARU, Palomar & CANDELS. Les distances sont calculées en supposant Λ CDM plat avec $H_0 = 70 \text{ km.s}^{-1}/\text{Mpc}$ et $(\Omega_\Lambda, \Omega_m) = (0.75, 0.25)$

Ainsi, moins de 10 SNe Ia explosent dans une sphère de $z = 0.01$ par an, près de 10 000 à $z = 0.1$ et jusqu’à 100 millions si le rayon est $z = 1$. Le nombre de SNe détectables en pratique est inférieur, car un instrument ne peut pas observer l’ensemble du ciel en continu pendant un an.

2.2.3 Explosion par effondrement gravitationnel

Nous nous intéressons aux SNe de type Ib, Ic et II.

Mécanisme explosion L’explosion est engendrée par une étoile massive arrivant en fin de vie. Après avoir fusionné tous les éléments légers composant son cœur, l’étoile ne produit plus assez d’énergie pour contre-balancer la force gravitationnelle et finit par s’effondrer sur elle-même. Cet effondrement s’arrête lorsque l’étoile est tellement compacte que c’est la force d’interaction forte – régissant la physique au cœur des noyaux et définissant une distance limite entre les particules élémentaires – qui stoppe l’effondrement. Cet arrêt provoque un rebond et une onde de choc brutale qui désintègre l’étoile. Les résidus de cette supernova sont soit une étoile à neutrons, pour les étoiles de faible masse, soit un trou noir pour celles de grande masse. Il existe une limite en masse à l’existence des étoiles à neutrons appelée masse d’Oppenheimer-Volkoff ($\sim 3,3 M_\odot$).

Diversité Pour ces SNe, il existe une grande diversité de progéniteurs. Les supernova de type Ib-c sont le résultat d’explosion d’étoiles très massives, de masses supérieures à $15 M_\odot$. Les supernovæ de type II proviennent d’étoiles massives, de masses supérieures à $8 M_\odot$. Ces SNe ont été observées seulement dans les galaxies formant des étoiles [Alsabti & Murdin (2017)].

On peut voir sur la figure 2.2, le spectre de quatre SNe appartenant à des sous-types différents. Les SNe de type Ib contiennent peu de silicium mais beaucoup d'hélium autour du maximum de luminosité. Les SN Ic sont définies par l'absence d'hélium et de silicium les spectres autour du maximum de luminosité, le spectre de SN2007gr (cf. Figure 2.2) ne contient ni hélium, ni silicium, ni hydrogène. Les supernovæ de type II peuvent être très facilement identifiées aux raies de Balmer $H\alpha$, $H\beta$ et $H\gamma$ de l'hydrogène présentes sur leurs spectres émis quelques semaines après le maximum. Ces raies sont visibles sur le spectre de la SN 1987A de la figure 2.2 à respectivement 6200, 4800 et 4100 Å.

On peut voir figure 2.3, les courbes de lumière, dans une bande bleue, des différents types de SNe. Les courbes de lumière des supernova de type b-c ont la même évolution. Les courbes de lumière des SN de type II se distinguent en deux grandes classes. Les SN de type II-P avec la stabilisation de la luminosité pendant quelques mois après le maximum d'explosion, tandis que les SN de type II-L ont une décroissance rapide. La luminosité de ces SNe est par ailleurs au minimum 10 fois moins importante que celle des SNe Ia.

2.2.4 Supernova thermonucléaire

Nous nous intéressons plus en détail aux SNe Ia. Elles sont caractérisées par l'absence d'hydrogène et par une forte proportion de silicium dans leur spectre (cf. Figure 2.7), avec une raie caractéristique à 6100 Å. En pratique, pour affirmer qu'une supernova est de type Ia, on recherche cette raie dans le spectre du candidat. Le spectre de SN Ia contient aussi du soufre, du calcium et plusieurs éléments du groupe fer.

Mécanisme explosion L'explosion thermonucléaire (SNe Ia) est produite par une naine blanche composée principalement de carbone et d'oxygène dont la masse augmente pour approcher une limite, la masse de Chandrasekhar, d'environ $1.44 M_{\odot}$ [Chandrasekhar (1931)], à proximité de laquelle l'étoile explose.

Une naine blanche est une étoile dense et dégénérée, ultime stade de l'évolution stellaire d'étoiles de masses inférieures à environ $8 M_{\odot}$. Le terme "dégénéré" signifie que la matière est dans un état suffisamment dense, pour que le principe d'exclusion de Pauli² se manifeste aux échelles macroscopiques. Les atomes sont tellement proches les uns des autres, que les électrons des couches externes ne peuvent plus se rapprocher davantage sans violer le principe d'exclusion de Pauli. Il se crée une *pression de dégénérescence* qui empêche une augmentation de la densité. Si la masse de la naine blanche augmente, la fusion du carbone commence, induisant une augmentation de la température.

Pour une étoile normale de la séquence principale du diagramme de Hertzsprung-Russell, l'augmentation de température entraîne une augmentation de volume, qui par suite refroidit l'étoile et stabilise le processus. Pour une naine blanche, la pression de dégénérescence n'étant que très faiblement couplée à la température, les conditions de pression nécessaires à la fusion du carbone et par suite de l'oxygène sont maintenues. La réaction devient alors une réaction thermonucléaire en chaîne, une fraction importante du carbone et de l'oxygène fusionne en une fraction de seconde [Röpke & Hillebrandt (2004)]. L'énergie libérée, environ $10^{51} J$ désintègre l'étoile, en expulsant la matière à des vitesses de l'ordre d'une dizaine de milliers de kilomètres par seconde, c'est une *Supernova de type Ia*. La majorité de l'énergie est convertie en énergie cinétique et seule une fraction d'environ $\sim 10^{44} J$ [Khokhlov et al. (1993)] est convertie en énergie lumineuse. Elle provient de la désintégration des éléments radioactifs créés lors de l'explosion et est responsable

2. Phénomène purement quantique interdisant à deux électrons d'être dans le même état quantique.

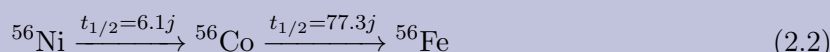
de la grande luminosité des supernovæ de type Ia. La luminosité d'une SN seule est du même ordre de grandeur que celle du coeur de sa galaxie hôte, soit environ $10^{10} L_{\odot}$ (cf. Figure 2.1b).

Ce scénario est corroboré par plusieurs observations. D'abord, l'observation de SN 2011fe [Bloom et al. (2012)], dont on peut voir les spectres sur la figure 2.6a, a démontré pour la première fois que l'étoile responsable de cette explosion était une étoile dégénérée et compacte sûrement une naine blanche. De plus, le fait que le phénomène se produise dans des conditions quasi-identiques (composition chimique, masse limite) et ne fait intervenir que des paramètres fondamentaux explique la similitude des luminosités observées. L'absence d'hydrogène et d'hélium dans les spectres des SNe peut s'expliquer par l'absence de ces éléments au moment de l'explosion et la présence de silicium, soufre et calcium par le fait que ce sont les produits de la fusion du carbone et de l'oxygène. Enfin que les SNe Ia soient les seules à avoir été observées dans des environnement d'étoiles vieilles, comme les galaxies elliptiques, se justifie par le fait que les naines blanches carbone-oxygène soient le dernier stade de l'évolution stellaire d'étoile de taille standard, trop froides pour avoir des réactions nucléaires et rayonnant leur énergie doucement pendant des millions d'années.

Progéniteurs Les naines blanches sont des étoiles trop peu lumineuses pour être observées directement à grande distance. Elles sont détectées peu après leur explosion. Ainsi le progéniteur d'une supernova Ia n'a jamais été observé. Les informations que nous avons sur les étoiles engendrant des SNe Ia sont seulement déduites indirectement. Il existe plusieurs scénarios pour expliquer l'explosion en SN Ia. Nous en présentons deux principaux :

- *Single degenerate* [Nomoto (1982)] : ce scénario suppose qu'une naine blanche de faible masse accrète de la masse d'une étoile compagne, de masse plus importante. Cette étoile doit donc avoir une gravité à la surface assez faible et être non dégénérée, comme une géante rouge ou une étoile en fin de séquence principale. Le matériel accrété, principalement de l'hydrogène, est fusionné en carbone puis en oxygène à la surface de l'étoile. L'explosion se produit lorsque la naine blanche approche la masse de Chandrasekhar. Ce scénario montre cependant quelques problèmes. Une accrétion trop lente, inférieure à $10^{-8} - 10^{-7} M_{\odot} \cdot \text{an}^{-1}$, ne pourrait pas compenser les mécanismes de perte de masse comme les flashes d'hélium ou les éruptions nova. Par contre une accrétion trop rapide, transformerait l'étoile en géante rouge, engloutissant l'étoile compagne et donc arrêtant l'accrétion. Le principal argument en défaveur de ce scénario est le fait que la matière accrétée est composée essentiellement d'hydrogène et d'hélium ainsi une accrétion de l'étoile compagne impliquerait de l'hydrogène dans le spectre des SNe Ia. Ce simple mécanisme est néanmoins favorisé par la reproductibilité de l'état initial, expliquant très simplement, la similitude des luminosités observées.
- *Double degenerate* [Iben & Tutukov (1984)] : ce scénario suppose l'échange de masse entre une naine blanche carbone-oxygène et une autre naine blanche compagne, orbitant en orbite proche, par fusion directe entre les deux objets ou bien par accrétion lente de la matière de la naine blanche la plus légère par effet de marées. Contrairement au scénario précédent, celui-ci explique simplement l'absence d'hydrogène dans le spectre des SNe Ia et il pourrait aussi expliquer l'observation d'événements extrêmement lumineux.

Source d'énergie Les SNe Ia puisent leur énergie de la désintégration des éléments du groupe du Fer qui ont été produits lors de l'explosion [Colgate & McKee (1969)]. Le principal élément radioactif formé est le ^{56}Ni qui se désintègre – avec une demi-vie de 6.1 jours – en ^{56}Co , qui lui-même se désintègre – avec une demi-vie de 77,3 jours – en ^{56}Fe stable :



Entre 0.3 et 0.8 M_{\odot} de Nickel sont produits lors de l'explosion d'une SN Ia. Avant le maximum de luminosité, la majorité de l'énergie lumineuse est émise sous forme de rayonnement gamma et sous forme de positrons. Ils sont ensuite absorbés par l'éjecta dense, puis ré-émis sous forme de rayonnement optique et proche infrarouge. Le rayonnement observé se situe pour plus de 85 % dans le visible. Le maximum de luminosité est supposé être proportionnel à la quantité de ${}^{56}\text{Ni}$ [Arnett (1982)]. Le ${}^{56}\text{Ni}$ étant le résultat de la fusion de la chaîne commençant avec le carbone et l'oxygène, c'est un argument supplémentaire en faveur des naines blanches carbone-oxygène comme progéniteur des SNe Ia.

2.3 Propriétés observationnelles des SNe Ia

Puisque la majorité du flux des SN Ia est émis dans le visible, c'est dans cette gamme de longueur d'onde que se font la majorité des observations. Nous allons détailler les caractéristiques des observations spectroscopiques et photométriques. On peut voir sur la figure 2.6a, l'évolution de spectres des SNe Ia en fonction des dates d'observations. On voit sur la figure 2.6b les courbes de lumière dans 8 bandes allant de l'ultraviolet u à l'infrarouge H .

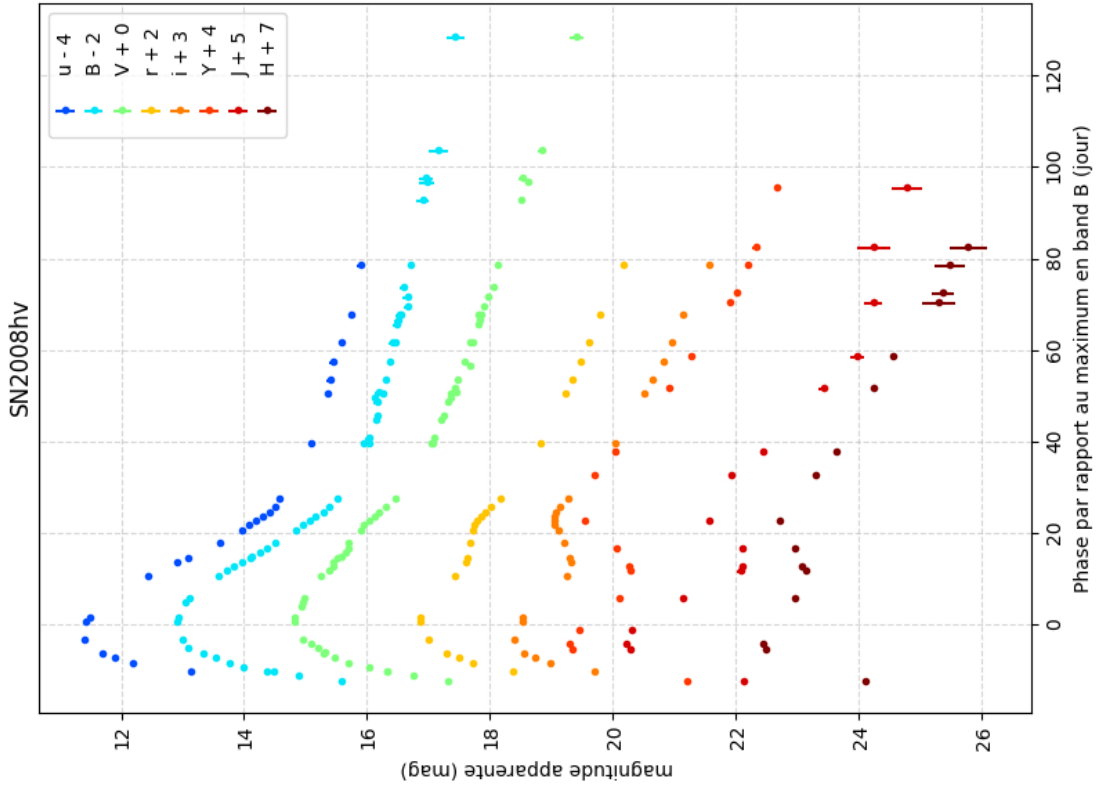
2.3.1 Propriétés spectrales

Sur la figure 2.6a, on peut voir un suivi spectroscopique d'une SN Ia, de 32 spectres, s'étalant sur plus de 100 jours. On peut voir sur la Figure 2.7, un spectre d'une SN Ia observé proche du maximum de luminosité en bande B.

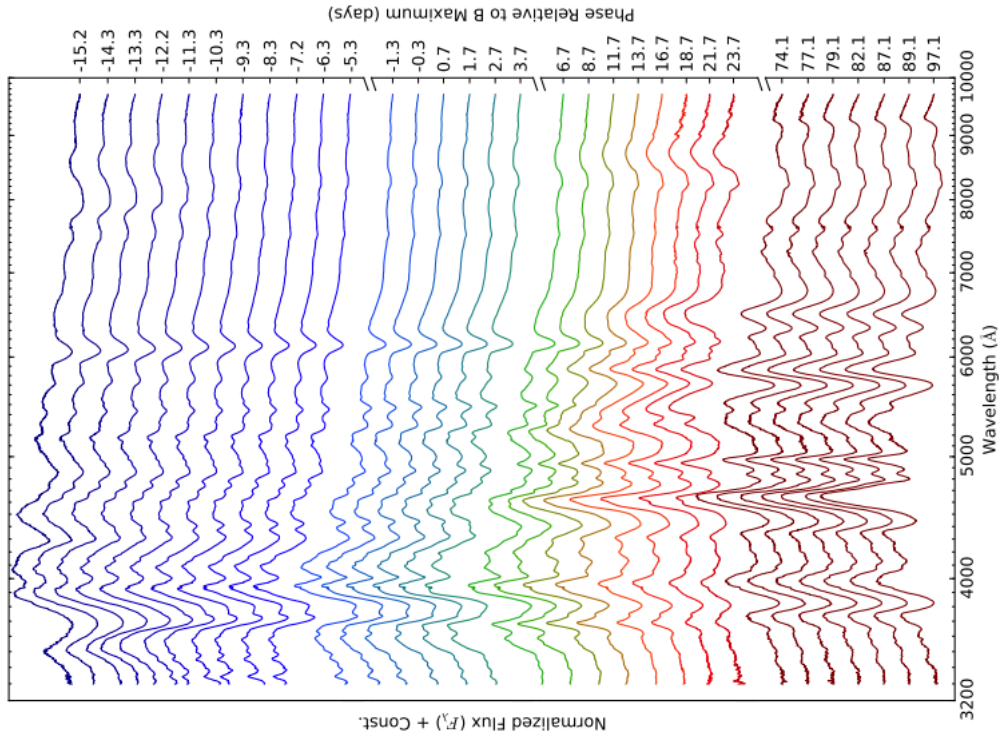
Du fait de la grande vitesse de l'éjecta ($\sim 10\,000\text{ km}\cdot\text{s}^{-1}$), les spectres de SNe Ia présentent des raies larges³. Une raie typique, présente aux alentours de 6100Å , est due au doublet du Si II à 6347 Å et 6371 Å . D'autres raies sont dues au silicium comme on peut le voir sur la figure. On voit aussi la présence dans l'éjecta de calcium à petite et grande longueur d'onde, de magnésium ou de fer. Non représentées sur la figure, mais présentes et augmentant avec la phase, il y a aussi des raies de nickel et de cobalt.

La forme des raies spectrales est caractéristique des enveloppes en expansion rapide, appelées *profil P-Cygni*. Dans ce modèle, la photosphère en expansion peut être divisée en trois régions (cf. Figure 2.8). Une zone de diffusion en gris clair, une autre région de diffusion en gris et une région occultée en gris foncé. Les photons de la photosphère, traversant la zone de diffusion en gris sur la ligne de visée, vont être décalés vers le bleu puisque l'éjecta s'approche de nous. De plus, ils sont diffusés dans toutes les directions, ce qui implique un déficit de flux aux longueurs d'onde plus petites que la longueur d'émission (cf. Figure 2.8). Les photons traversant la zone en gris clair vont eux aussi être diffusés, dont certains dans notre direction impliquant un excédent de flux, à la longueur d'onde d'émission s'ajoutant à la contribution de la photosphère. Ils sont équitablement décalés vers le rouge ou le bleu, centrant l'oscillation en λ_0 . La contribution de la région en gris foncé est négligeable. Les profils P-cygni dépendent de la vitesse d'expansion de la couche de matière, de la profondeur optique et de sa densité. Les profils P-cygni que nous voyons sur les spectres des SNe Ia, sont mélangés entre les différents éléments, les rendant moins distincts que ceux de la figure 2.8.

3. Impliquant qu'une haute résolution spectrale n'est pas indispensable pour les caractériser.



(b) Huites courbes de lumière de la SN2008hv allant de l'ultraviolet au proche infrarouge ($z \sim 0.014$).



(a) 32 spectres de SN2011fe ($z \sim 0.0012$). Figure extraite de Pereira et al. (2013).

FIGURE 2.6 – Suivi spectroscopique de SN2011fe et photométrique de SN2008hv. La proximité des objets permet un échantillonnage dense d'observations.

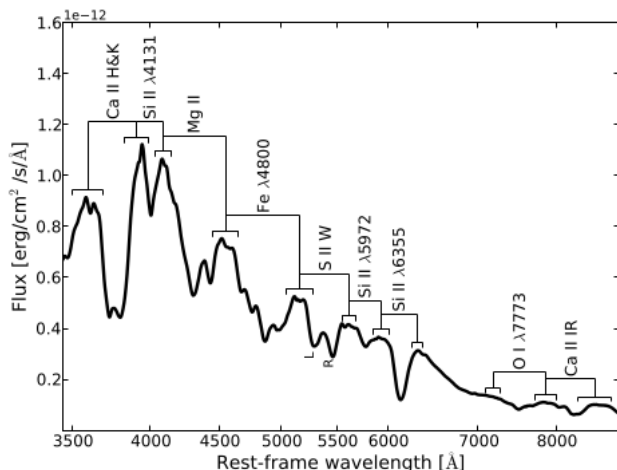


FIGURE 2.7 – Spectre typique d'une supernova au maximum de luminosité en bande B. Figure extraite de Chotard (2011).

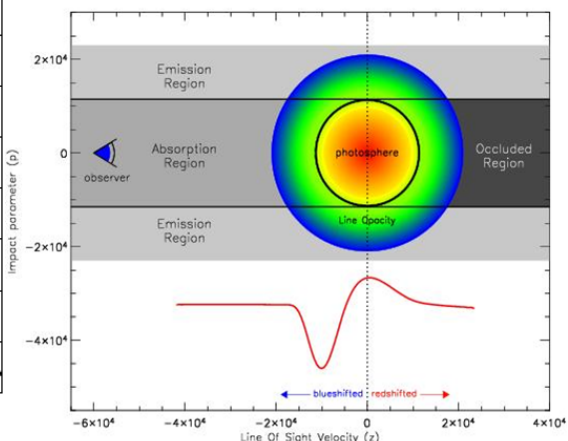


FIGURE 2.8 – Schématisation de profils P-Cygni, figure extraite de Roth & Kasen (2018).

On peut diviser l'évolution temporelle des spectres des SNe Ia en deux phases (cf. Figure 2.6a). La première phase, jusqu'à environ quelques jours après le maximum, les spectres montrent la présence des éléments typiques des SNe Ia avec de grandes vitesses d'éjecta. Autour du maximum de luminosité, on voit un rétrécissement de la largeur des raies, donc une diminution de la vitesse d'explosion autour de $10\,000\text{ km}\cdot\text{s}^{-1}$. Les spectres y sont homogènes, avec principalement les raies du silicium et du calcium. Pendant cette première phase, les couches externes d'éjecta sont opaques au rayonnement. L'opacité des éjectas devenant de moins en moins importante, une deuxième phase apparaît, pendant laquelle nous pouvons sonder les couches internes de la SN et d'observer des raies d'émissions spécifiques, notamment les émissions du cobalt, du nickel et du fer.

Une telle série d'observations spectroscopiques est très rare et se cantonne à des SNe de petits *redshifts*⁴. Les spectres des supernovæ de type Ia sont généralement observés pour caractériser l'objet en se référant aux raies typiques et obtenir un *redshift*. En pratique, la grande statistique d'observations est obtenue en photométrie.

2.3.2 Propriétés photométriques

Les SNe Ia ont une luminosité qui augmente pendant une vingtaine de jours pour atteindre un maximum, dont la magnitude absolue, en bande *B*, est de $-19,05$ mag environ. On choisit comme référence de temps, la date de maximum en bande *B* dans le référentiel de la supernova par convention, qu'on note *p*.

Visible Les SNe Ia ont une luminosité très importante, environ $10^{10} L_{\odot}$, ce qui nous permet de les voir à grand *redshift*, jusqu'à $z \sim 1.5 - 2$. Après la phase de fusion, c'est la chaîne de désintégration, définie équation 2.2 qui est la source de luminosité de la SN. L'évolution de la courbe de lumière est déterminée par la quantité de ^{56}Ni produite lors de la fusion du carbone et de l'oxygène présents dans le coeur de la naine blanche. La courbe de lumière en bande *B* (cf. Figure 2.6b) présente une première partie, entre $p = -10j$ et $p = 25j$, pour laquelle la luminosité provient de la désintégration du ^{56}Ni en ^{56}Co . A cette phase, le flux reçu de la SN

4. SN2011fe est la SN la plus proche de la Terre jamais observée.

n'est pas maximum, puisque la matière est encore dense et opaque. Au fur et à mesure que le temps s'écoule, l'énergie lumineuse émise diminue mais la taille de l'éjecta augmente, diminuant la densité et donc l'opacité de la matière laissant échapper plus de photons [Arnett (1982)]. Ainsi l'énergie lumineuse pouvant s'extraire de l'enveloppe augmente. La première partie croissante de la courbe de lumière est due à ce phénomène et est appelée "temps de montée". Il est typiquement en bande B d'environ 20 jours. Au-delà de 25 jours, on aperçoit une décroissance exponentielle suivant le temps de demi-vie caractéristique de la chaîne de désintégration du ^{56}Co en ^{56}Fe . Après 50 jours, la courbe de lumière suit un déclin exponentiel et perd une magnitude par mois environ.

Proche infrarouge Dans cette gamme correspondant aux bandes r, i, I, Y, J, H, K , vers 25 jours après le maximum de luminosité en bande B , les SNe Ia présentent un second pic de luminosité, dû à la désexcitation du Fer III en Fer II. Elle augmente la transparence dans ces longueurs d'ondes et se produit lorsque la SN se refroidit. À ce moment, la luminosité dans cette gamme de longueur d'onde est à son maximum, environ 20 % de la luminosité totale de la SN [Wang et al. (2009)]. Les courbes de lumière des bandes infrarouges sont remarquablement homogènes, avec par exemple une variation de 0.09 magnitude en bande H . Cette faible dispersion s'explique notamment par une absorption des poussières présentes sur la ligne de visée, plus faible que dans les bandes du visible. Les supernovæ sont dans ces bandes de longueurs d'onde, des chandelles standard, présentant seulement 6% d'incertitude sur la distance en utilisant des courbes de lumière en bande H [Barone-Nugent et al. (2012)], sans utilisation de paramètres de standardisation. Cependant ces observations sont bien plus complexes à obtenir depuis le sol puisque le rayonnement infra-rouge est absorbé par l'atmosphère.

Ultraviolet Dans cette gamme de longueur d'onde (bande u , inférieur à 4000 Å) la proportion de flux émis est bien plus faible que dans le visible. Le maximum de flux dans l'UV est de 10 % du flux total [Wang et al. (2009)]. Les courbes de lumière y suivent la même évolution que dans le visible. Les spectres observés dans l'UV sont plus sensibles aux métaux de l'éjecta. Ceci explique une dispersion plus grande des courbes de lumière observées en UV que dans celles du visible. Les observations montrent une augmentation de la luminosité dans cette gamme de longueur d'onde avec le *redshift* [Foley et al. (2012)].

2.3.3 Standardisation

Même si les courbes de lumière des bandes visibles des supernovæ Ia sont très homogènes, elles présentent néanmoins quelques variations. Les SNe ne sont donc pas des chandelles standards mais bien des chandelles standardisables. Historiquement, l'utilisation des SNe comme indicateur de distance repose sur le fait que leur magnitude absolue en bande B peut être standardisée en utilisant des propriétés photométriques observationnelles. En effet, si on ne tient pas compte de la standardisation, on observe une dispersion au pic de luminosité en bande B de 40 %. La relation de Tripp [Tripp & Branch (1999)] utilisant deux paramètres de standardisation, un paramètre de forme et un de couleur, permet de diminuer la dispersion au pic à environ 15%. Avant corrections, la magnitude absolue au maximum de luminosité en bande B est comprise entre -18 et -19.5 mag, avec une valeur de 19.50 ± 0.03 dans le système de magnitude AB [Fukugita et al. (1996)] avec $H_0 = 68.50 \pm 1.2$, lorsque qu'une cosmologie $w - \Lambda\text{CDM}$ est ajusté en combinant les résultats de CMB+BAO+SNe, selon Betoule et al. (2014).

Brighter-slower relation

Cette première relation lie le maximum de luminosité en bande B à l'évolution temporelle de la courbe de lumière. Il a été observé [Rust (1974), Pskovskii (1977), Phillips (1993)] que plus une supernova est lumineuse plus son évolution est lente.

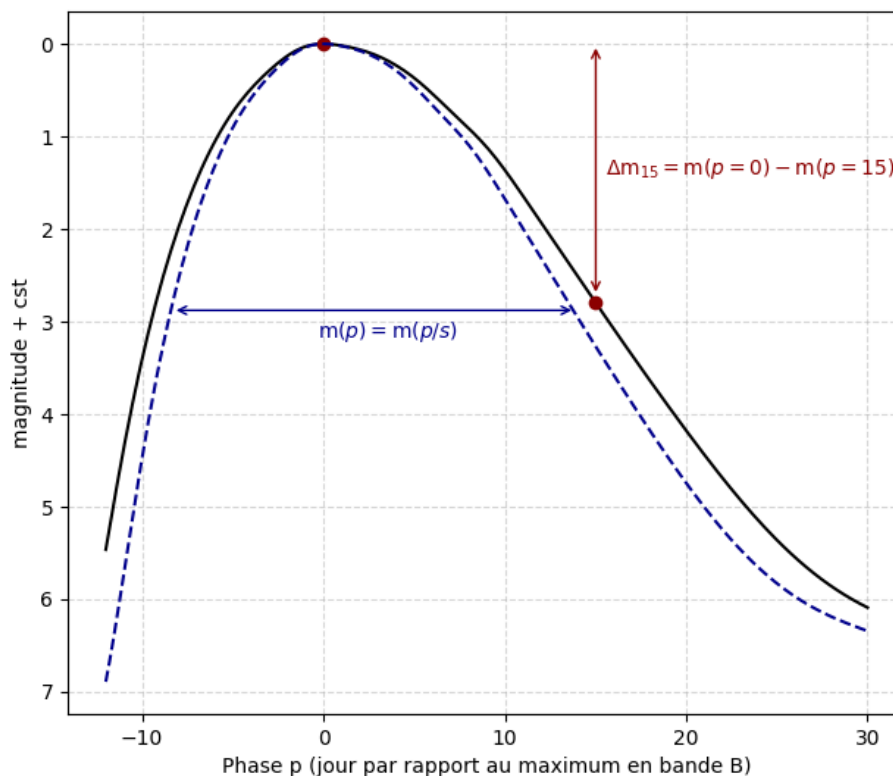


FIGURE 2.9 – Schéma de deux paramétrisations de l'évolution temporelle de courbes de lumière en bande B avec le paramètre Δm_{15} en rouge et le stretch, s en bleu. Le modèle en noir peut soit être paramétrisé selon son écart en magnitude entre sa valeur à la phase 0 et +15 jours, soit défini selon le *stretch*, dilatation temporelle des phases.

Le premier estimateur quantifiant l'évolution temporelle de la classe proposé par Phillips (1993) est la différence de magnitude absolue en bande B , entre le maximum de luminosité et 15 jours après, noté Δm_{15} :

$$\Delta m_{15} = M_B(t = t_{max}^B) - M_B(t = t_{max}^B + 15) \quad (2.3)$$

$$\Delta m_{15} = M_B(p = 0) - M_B(p = 15) \quad (2.4)$$

Depuis, d'autres estimateurs ont été utilisés, pour cette même relation prenant en compte l'information de l'évolution dans d'autres bandes ou encore une normalisation de l'amplitude de la courbe de lumière. Une paramétrisation possible est le *stretch* défini comme une dilatation temporelle de la courbe de lumière M_B par rapport au modèle moyen M_B^{model} dans ce cas et noté s :

$$M_B(p) = M_B^{model} \left(\frac{p}{s} \right) \quad (2.5)$$

Pour une SN plus lumineuse, son temps d'évolution sera plus lent, signifiant un paramètre $s > 1$. Ces deux paramétrisations sont illustrées sur la figure 2.9. Ensuite, [Guy et al. \(2007\)](#) introduit un paramètre X_1 relié au *stretch*, tenant compte de la variabilité de forme des courbes de lumière. Sur la figure 2.10a, on peut voir la relation entre la magnitude absolue au maximum de luminosité en bande B et x_1 , pour les 740 supernovæ de l'analyse JLA [[Betoule et al. \(2014\)](#)]. Cette variabilité est considérée comme intrinsèque et liée à la quantité de Nickel 56 formé lors de la fusion. La largeur de la courbe de lumière est déterminée par le temps de diffusion d'un photon dans l'éjecta qui dépend de l'opacité et de la quantité de d'énergie lumineuse créée [[Arnett \(1982\)](#)].

Brighter-bluer relation

La seconde source de variabilité a été observée par [Hamuy et al. \(1995\)](#) et [Riess et al. \(1996\)](#). Elle lie la luminosité des SNe à leur couleur. Ce phénomène serait provoqué par des effets extrinsèques et intrinsèques. L'effet extrinsèque provient de l'absorption de la lumière sur la ligne de visée par la poussière de la galaxie hôte et du milieu interstellaire. Cette absorption est un phénomène chromatique. Ainsi les photons de faibles longueurs d'onde seront plus facilement absorbés, impliquant l'observation d'une SN rougie. Plus la lumière traversera de poussière plus elle sera rougie. L'effet intrinsèque est lié aux propriétés physiques du progéniteur. [Phillips \(1993\)](#) paramétrise la couleur d'une SN donnée par la différence de sa couleur au maximum et de la couleur moyenne des SNe au maximum :

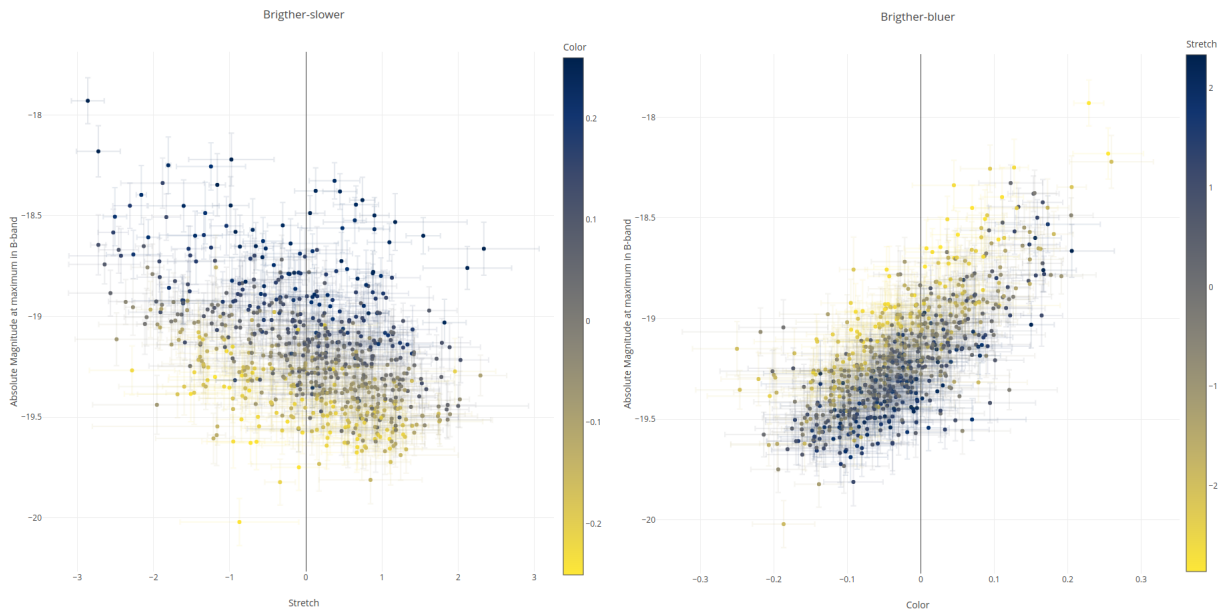
$$c = (B - V)_{max} - \langle B - V \rangle_{max} \quad (2.6)$$

Sur la figure 2.10b, on peut voir que plus une supernova est lumineuse, c'est à dire plus sa magnitude absolue dans la bande B au maximum de luminosité est petite, plus sa couleur a tendance à être négative, c'est-à-dire bleue.

Relation de standardisation On peut donc diminuer la dispersion sur la modélisation de magnitude absolue en bande B des SNe observées \tilde{M}_B , donc des résidus au diagramme de Hubble, en combinant les deux relations de *brighter-slower* et *brighter-bluer*, en utilisant la magnitude absolue standardisée M_B [[Tripp & Branch \(1999\)](#)] :

$$\begin{aligned} M_B &= \tilde{M}_B + 40\% \\ M_B &= \tilde{M}_B - \alpha X_1 + \beta c + 15\% \end{aligned} \quad (2.7)$$

M_B est la magnitude absolue en bande B, X_1 est le paramètre de forme tenant compte de la relation de *brighter-slower* et c le paramètre de couleur tenant compte de la corrélation *brighter-bluer*. Les paramètres \tilde{M}_B , α et β sont estimés sur l'ensemble du lot de SNe utilisées pour la mesure des paramètres cosmologiques et sont donc des paramètres de nuisance. X_1 et c doivent être extraits des données. La physique de ces objets étant complexe, les progéniteurs et le mécanisme d'explosions doivent être d'abord mieux compris dans le but d'avoir des simulations réalistes de l'explosion des SNe. Pour mesurer des distances et reconstruire le flux des SNe, on utilise des modèles empiriques qui extraient les informations de standardisation des SNe directement des



(a) Relation liant l'étirement temporel des supernovæ à leur luminosité absolue.

(b) Relation liant la couleur des supernovæ à leur luminosité absolue.

FIGURE 2.10 – Loi de standardisation de 740 SNe Ia extraite de la Joint Light Curve Anlysis (JLA) [Betoule et al. \(2014\)](#).

données. La création et la compréhension de ces modèles empiriques de flux des SNe est le coeur du travail réalisé au cours de cette thèse.

2.3.4 Propriétés de la galaxie hôte

Les galaxies dans lesquelles se produisent des SN Ia évoluent avec le *redshift*. La nature de ces galaxies a un impact sur la luminosité des SNe Ia : les galaxies spirales et irrégulières produisent des SN Ia en moyenne plus lumineuses que les SNe venant de galaxies elliptiques ou lenticulaires [[Hamuy et al. \(1996\)](#)]. Cette dépendance n'a pas d'impact sur les résultats cosmologiques car elle est correctement corrigée par la standardisation avec les relations de *brigther-slower*. Ce n'est pas un effet d'extinction car les galaxies spirales possèdent des populations stellaires plus jeunes que les galaxies elliptiques, donc avec plus de poussière, ce qui induirait plus d'extinction et non l'inverse. Cette dépendance de la luminosité des SNe Ia avec le type de la galaxie hôte, dont la distribution varie avec le *redshift*, a été observé dans le lot de SNe de SNLS [[Howell et al. \(2007\)](#)].

Avec l'agrandissement des lots de SNe, une dépendance entre la luminosité standardisée M_B et les propriétés de la galaxie hôte a finalement été observée : avec la masse des galaxies [[Kelly et al. \(2010\)](#), [Sullivan et al. \(2010\)](#)], ou encore le taux de formation stellaire des galaxies hôtes [[Lampeitl et al. \(2010\)](#)]. Cette observation motiva une division de la masse stellaire de la galaxie hôte en deux intervalles, inférieur et supérieur à $10^{10} M_{\odot}$. L'ajustement d'une loi indépendante pour chacun de ces deux lots permit d'obtenir une réduction de la dispersion [[Childress et al. \(2013\)](#)].

La luminosité standardisée M_B des SNe de JLA appartenant à des galaxies de masse stellaire supérieure à $10^{10} M_{\odot}$ est supérieur de 0.06 mag à celle appartenant à des galaxies de masse inférieure à cette limite (cf. Figure 2.11).

On peut donc ré-écrire l'équation 2.7 :

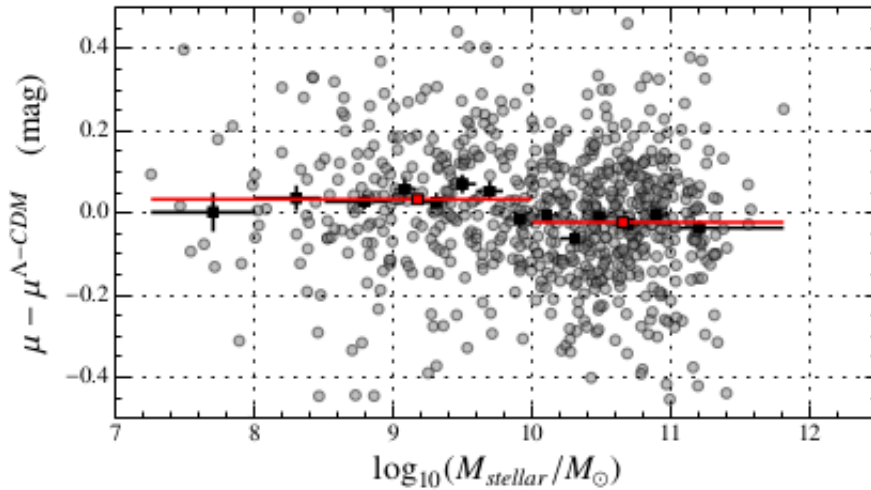


FIGURE 2.11 – Résidus au diagramme de Hubble de l’analyse JLA [Betoule et al. (2014)] en fonction du logarithme de la masse stellaire de la galaxie hôte. On constate un écart de 0.06 mag entre les deux populations de SNe défini par $M_{\text{stellaire}} < 10^{10} M_{\odot}$ et $M_{\text{stellaire}} \geq M_{\odot}$.

$$M_B = -\alpha X_1 + \beta c + \begin{cases} \tilde{M}_B & \text{si } M_{\text{stellaire}} \leq 10^{10} M_{\odot} \\ \tilde{M}_B + \Delta_M & \text{sinon} \end{cases} \quad (2.8)$$

$M_{\text{stellaire}}$ étant la masse stellaire de l’hôte et Δ_M un quatrième paramètre de nuisance, d’autres caractéristiques de la SNe que la masse de sa galaxie hôte peuvent être utilisées. Ainsi une meilleure corrélation entre la luminosité au maximum et l’âge de la population stellaire de l’environnement local de la SN a été établi, en mesurant la couleur de la galaxie hôte dans un rayon de 3 kpc autour de la SN [Roman et al. (2018)] ou la luminosité de la raie H α dans un rayon de 1 kpc autour de la SN [Rigault et al. (2020)]. Du fait de l’évolution de la population des galaxies avec le *redshift*, ces caractéristiques varient elles aussi en fonction du *redshift*. Il est donc crucial de tenir compte de cet effet afin de ne pas induire de biais dans l’estimation des paramètres cosmologiques.

2.3.5 Sous-classes

Si la vaste majorité des SNe Ia est standardisable, un tiers des SNe observées montrent des caractéristiques différentes. Ces supernovæ, dites *particulières*, ont été classifiées en sous types (cf. Figure 2.13) aux frontières perméables. Sur la figure 2.13, sont représentées les magnitudes absolues non standardisées en bande B des différentes SNe Ia appartenant aux sous-classes en fonction du paramètre Δm_{15} des individus. On observe que ces SNe Ia particulières, présentant pourtant les caractéristiques spectrales des SNs Ia sont soit plus lumineuses, soit moins lumineuses que la moyenne, même en prenant en compte la correction de *brighter-slower*. Leur nom provient de la première SN Ia observée appartenant à la classe. Nous allons voir les caractéristiques des trois sous-types, les plus abondants. Le diagramme suivant présente la part des quatre grandes classes des SNe Ia observées dans un volume limité [Li et al. (2011)].

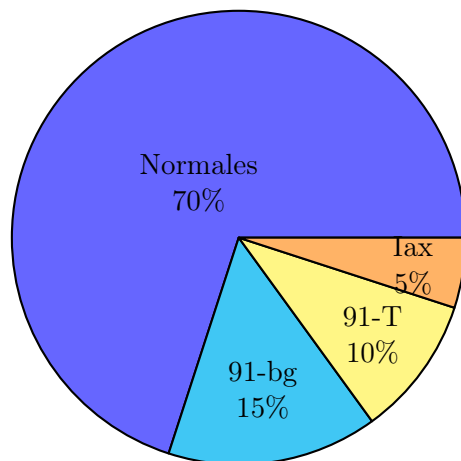


FIGURE 2.12 – Proportions des sous-classes de SNe Ia dans un volume fini [Li et al. (2011)].

Super-Lumineuse 1991T [Filippenko et al. (1992a)] - est une classe sur-lumineuse de SN Ia caractérisée par une grande luminosité au maximum. Blondin et al. (2012) montre que les SNe appartenant à cette classe sont en moyenne 0.2 - 0.3 mag plus lumineuses que les SNe standards et montrent aussi un Δm_{15} plus petit, correspondant à un stretch plus grand (cf. Figure 2.13). Ces objets montrent un second pic dans l'infrarouge plus important que les SNe Ia *normales*. Cette classe est caractérisée par l'absence de la seconde raie de Si II dans le spectre au maximum.

Sous-lumineuse 1991bg [Filippenko et al. (1992b)] - est une classe sous-lumineuse, caractérisée par une luminosité plus faible que celle des SNe Ia *normales* et qui se distingue par l'absence de second pic de luminosité dans le proche infra rouge, une caractéristique essentielle des SNe Ia. Taubenberger et al. (2008) montre que la luminosité en bande B de ces objets au maximum de luminosité est inférieur de 1.5 à 2.5 mag dans les bandes visibles que les SNe *normales*. En bande B , elles sont environ 10 fois moins lumineuses. De plus, leur couleur est au maximum plus rouge que celle des SNe Ia *normales* : $(B - V)_{max}$ environ 0.5 - 0.6 mag. Elles montrent un rapide déclin de luminosité après le maximum Δm_{15} 1.8 - 2.1 (cf. Figure 2.13).

Sous-lumineuse 2002cx ou SN Iax [Phillips et al. (2007) - Foley et al. (2009)]- est aussi une classe sous lumineuse. Leur magnitude absolue en bande B , est comprise entre -19 et -13 pour les moins brillantes (cf. Figure 2.13). Cette grande classe regroupe plus de 50 objets particuliers, qui présentent tous des spectres plus lisses et de petites vitesses d'éjecta. Leur courbes de lumière ont des temps de montée plus faible et ne présentent pas de second pic dans le rouge, proche infra-rouge.

Cette courte présentation des SN Ia particulières donne une vue, non exhaustive, de la zoologie

de ces objets complexes mais leur étude n'est pas l'objet de ce manuscrit. Dans le but d'étudier, la modélisation de l'évolution spectrale des SNe Ia, il sera utile de garder à l'esprit qu'un lot de données peut être contaminé par certaines SNe Ia non standards et non standardisables.

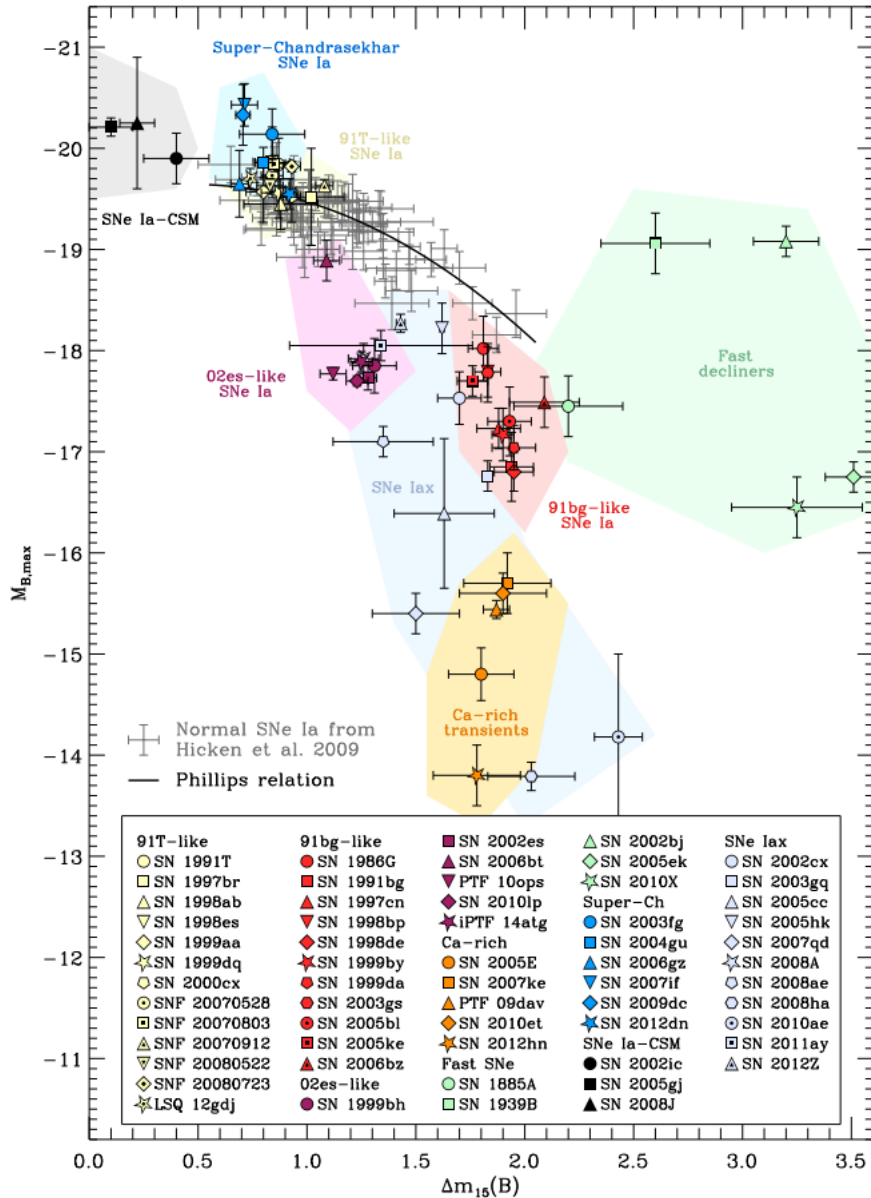


FIGURE 2.13 – Ce graphique, extrait de Taubenberger (2017), montre la diversité des SN Ia. Il représente l'évolution du paramètre de forme Δm_{15} en fonction de la magnitude absolue de représentants des différentes sous-classes des SNe Ia.

2.4 Conclusion

En somme, nous avons vu que les SNe Ia sont des objets complexes, avec des caractéristiques nous permettant non seulement de les observer à de grandes distances mais aussi de standardiser leur luminosité pour pouvoir mesurer ces distances précisément. Un de ces trois paramètres dépend de la physique de la galaxie hôte ce qui peut laisser penser que des paramètres de standardisation supplémentaires pourraient permettre de diminuer davantage la dispersion de leur luminosité.

La physique des objets qui les engendrent n'est pas le sujet de ce manuscrit mais l'amélioration de cette connaissance serait utile pour améliorer la précision de la mesure de la cosmologie. Néanmoins, pour l'instant les modèles de flux ne sont pas assez quantitatifs pour pouvoir être utilisés dans une analyse cosmologique. L'extraction des paramètres de standardisation dans le but de mesurer des distances nécessite donc une approche empirique et statistique. Dans le chapitre suivant, nous décrivons comment les observations de SNe sont réalisées et comment les contraintes cosmologiques en sont extraites.

Observation de SNe pour une analyse cosmologique

Les supernovæ de type Ia peuvent donc être utilisées pour contraindre l'équation d'état de l'énergie noire, en ajustant les paramètres cosmologiques impliqués dans le calcul des distances de luminosité. La technique est statistiquement très puissante : une variation de 10% (resp 5%) du paramètre w se traduit par un effet de l'ordre de 2% (resp 1%) à $z \sim 0.1$ et du même ordre de grandeur à $z \sim 1$. Les supernovæ présentant une dispersion résiduelle de 15% après standardisation, une centaine d'objets seulement permet de contraindre w à 10%, et un millier d'objets suffit pour atteindre une incertitude statistique de 5%. Dans ce chapitre, nous présentons les principales étapes des traitements qui permettent d'estimer des distances de luminosité standardisées à partir des images. Nous présentons tout d'abord les avancées technologiques qui ont permis de réaliser des mesures astronomiques de précision. Nous décrivons ensuite les techniques de suivi photométrique et spectroscopique les plus courantes. La section 3.3 présente les principaux relevés à l'origine des lots de données utilisées dans les diagrammes de Hubble récents. Dans la partie suivante, nous décrivons l'analyse du diagramme de Hubble. La section 3.5 décrit les sources d'incertitudes de cette analyse.

Contents

3.1	Les instruments de mesure	71
3.1.1	Une première révolution : le capteur CCD	71
3.1.2	Une seconde révolution : les caméras grand champ	73
3.1.3	Observations spectroscopiques	75
3.2	Données de suivi	77
3.2.1	Photométrie en bande large	78
3.2.2	Analyse des données photométriques	79
3.2.3	Données spectroscopiques	83
3.3	Les relevés	84
3.3.1	Première génération	84
3.3.2	Génération 1.5	86
3.3.3	Seconde génération	87
3.3.4	Troisième génération	88
3.3.5	Futurs relevés	89
3.4	Diagramme de Hubble - contraintes sur l'équation d'état de l'énergie noire	90
3.4.1	Mesure de distance avec des SNe Ia	90
3.4.2	Modèle empirique d'ajustement de courbes de lumière	91
3.4.3	Diagramme de Hubble	91
3.5	Sources d'incertitudes	92
3.5.1	Incertainitudes statistiques	93
3.5.2	Incertainitudes systématiques	94
3.6	Conclusion	96

3.1 Les instruments de mesure

Plusieurs avancées technologiques ont rendu possible l'observation à haute cadence de grands volumes d'Univers, permettant la constitution de grands échantillons de supernovæ. La première révolution concerne les détecteurs astronomiques.

3.1.1 Une première révolution : le capteur CCD

Le dispositif à couplage de charges, ou CCD (pour *Charge Coupled Device*) est un type de senseur utilisé en astronomie, depuis les années 1980. Un CCD est constitué d'un substrat semi-conducteur dopé n ou p , polarisé de façon à créer une zone de collection déplétée. Les photons entrant dans la zone de collection, s'ils sont d'énergie supérieure au *band-gap*¹ du semi-conducteur, peuvent créer une paire électron-trou par effet photo-électrique. Ceux-ci dérivent dans le champ de polarisation. Le champ de dérive est conçu de façon à former pour chaque pixel un puits de potentiel local dans lequel les photo-électrons créés durant la pose sont collectés. La lecture du CCD est destructive. Le principe est illustré sur la figure 3.1. Le CCD est équipé d'un registre de lecture, appelé registre série. Une fois la pose terminée, un dispositif d'horloges alternées permet de modifier les potentiels afin que les charges de chaque ligne soient transférées dans la ligne suivante, en direction du registre série, celui-ci collectant les charges de la ligne qui lui est adjacente. Les pixels du registre série sont ensuite lus par un ampli de charge. Le cycle de lecture reprend jusqu'à ce que le contenu de chaque pixel ait été collecté.

L'ampli de charge est composé d'une petite capacité (quelques fF) connectée à une chaîne d'amplification. Le signal amplifié est ensuite numérisé et stocké par le système d'acquisition en unités d'ADC ou ADU. La relation entre le niveau de signal numérisé et le nombre de photoélectrons collectés dans le pixel est du type :

$$\phi_{e^-} = g \times (\phi_{ADU} - p) \quad (3.1)$$

où g est le gain de la chaîne de lecture (en photo-électrons par unité d'ADC ou ADU) et p est un piédestal. Le processus de lecture introduit un bruit, dominé par le vidage de la capacité de lecture entre chaque pixel. Afin de limiter le bruit de lecture, la vitesse de lecture est maintenue assez lente, au niveau de quelques centaines de kilo-pixels par seconde. À température ambiante, l'agitation thermique induit la création spontanée de paires électrons-trous et donc l'apparition d'un courant d'obscurité, même en l'absence de lumière. Le phénomène n'est pas gênant dans les situations où le signal à mesurer excède largement le courant d'obscurité. En astronomie, domaine où les flux que l'on cherche à mesurer sont ténus et les temps de pose longs (de quelques dizaines à quelques milliers de secondes) le courant d'obscurité est limité en maintenant la température du CCD dans une plage de température généralement comprise entre -80°C et 0°C .

La création de charges dépend de l'énergie du photon incident donc de sa longueur d'onde. Il faut que l'énergie du photon soit suffisante pour faire passer un électron de la bande de valence à la bande de conduction. La fraction des photons incidents qui créent un photo-électron, est appelée *efficacité quantique* du CCD. L'efficacité quantique est fonction de la longueur d'onde. Elle est généralement notée $\epsilon(\lambda)$ ou $\eta(\lambda)$.

On présente figure 3.2 l'efficacité quantique de deux caméras : MegaCam, dont la conception remonte aux années 1999-2000, équipée de CCD minces ($40\mu\text{m}$), et Subaru/HyperSuprimeCam construite dix ans plus tard, et équipée de CCD haute-résistivité, plus épais ($200\mu\text{m}$). On constate que dans les deux cas, la sensibilité diminue dans l'infrarouge pour atteindre 0 vers 11000 \AA .

1. Énergie nécessaire pour qu'un électron du substrat passe de la bande de valence à la bande de conduction.

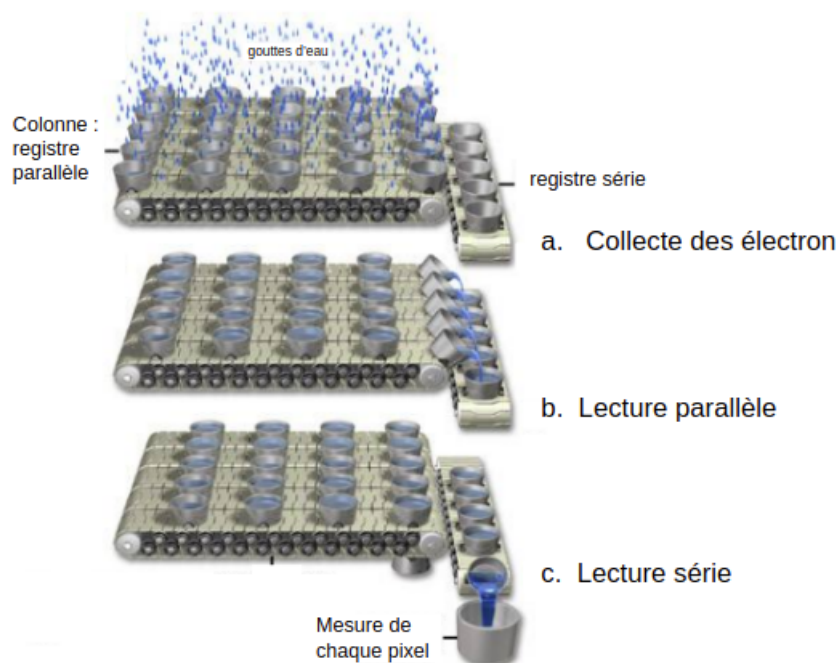


FIGURE 3.1 – Schéma de la collecte et du transfert de charge dans un CCD.

En effet, l'énergie minimale pour qu'un électron de la bande de valence rejoigne la bande de conduction est de 1.14 eV , correspondant à l'absorption d'un photon de $10,866 \text{ \AA} \approx 1 \mu\text{m}$. Les détecteurs silicium sont donc transparents aux longueurs d'onde plus élevées. En-deçà de $\lambda \sim 1 \mu\text{m}$, l'efficacité est fonction de l'épaisseur du senseur : à une température de -120°C , la longueur d'absorption d'un photon de 5220 \AA (respectivement 6350 \AA , 8500 \AA et 10000 \AA) est de $1.5 \mu\text{m}$ (respectivement 4 , 20 et $120 \mu\text{m}$). L'avènement des CCD haute résistivité a permis d'augmenter l'épaisseur des senseurs et de gagner significativement en efficacité quantique dans le rouge et l'infrarouge proche. Il existe par ailleurs des semi-conducteurs à faible écart entre les bandes de valence et conduction, le germanium par exemple, permettant d'aller plus loin dans l'infrarouge jusqu'à 16000 \AA .

La chute de sensibilité dans le bleu provient de plusieurs causes. Tout d'abord, l'indice du Silicium est $n \sim 4$, ce qui fait que les photons ont plus de chances d'être réfléchis à la surface du CCD. L'efficacité dans le bleu dépend donc crucialement de la qualité du *coating* anti-réfléchissant. Ensuite, la longueur d'absorption décroît fortement en fonction de l'énergie du photon. Les photons bleus ont une plus grande probabilité d'interaction sur une petite distance de pénétration. Ils ont donc une plus forte probabilité d'être absorbés hors de la zone de déplétion sans que le photo-électron soit capturé dans le puit de potentiel du pixel. Ainsi les CCD sont des instruments d'imagerie du visible, domaine dans lequel la majorité du flux des SNe Ia est émis (cf. section 2.3).

Une autre propriété importante des CCD est la linéarité de stockage des charges, qui permet de comparer quantitativement des poses longues et des poses courtes, et ainsi d'observer des objets peu lumineux simplement en augmentant la durée d'exposition.

Aujourd'hui le format standard est une matrice de 2048 par 2048 pixels. Néanmoins, une limitation importante des CCD est la difficulté à réaliser des surfaces homogènes sur plusieurs millions

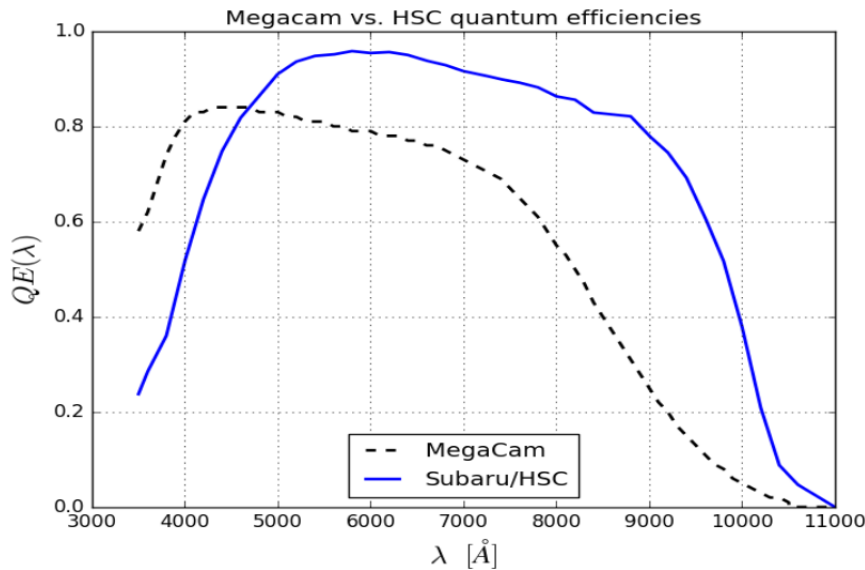


FIGURE 3.2 – Efficacité quantique de deux caméras, Mégacam utilisée pour les observations du relevé SNLS et HSC caméra du télescope Subaru (cf. Figure 3.13b). Les CCD de cette dernière sont optimisés pour observer dans le rouge.

de pixels.

3.1.2 Une seconde révolution : les caméras grand champ

Au tournant du millénaire, le développement de CCD aboutables a permis d’assembler de grandes mosaïques de senseurs, avec un espace inutilisable entre chaque CCD réduit à quelques dizaines de pixels (15 à 25 pixels [Léna (1996)]). Il est ainsi devenu possible de construire des plans focaux de plusieurs centaines de millions de pixels, et couvrant des surfaces de ciel de plusieurs degrés carrés avec une résolution spatiale de l’ordre de 0.2 secondes d’arc par pixel.

Plusieurs générations d’instruments grand champ ont été développés sur ce principe durant les trois dernières décennies. La figure 3.3 présente quelques exemples représentatifs, construits dans les années 1990-2000 (CFHT12k), 2000-2010 (DECam, HSC) et 2010+ (Observatoire Vera Rubin). La caméra CFHT12k est l’une des premières du genre. Elle a vu sa première lumière en 1999. Son plan focal est composé de 12 CCD pour plus de $1,05 \times 10^8$ pixels. Elle est montée au foyer primaire du télescope Canada France Hawaii (CFHT pour *Canada France Hawaii Telescope*, cf. Figure 3.13c), de 3.5m, et couvre 0.3 deg^2 [Starr et al. (2000)]. La caméra de génération suivante (MegaCam), mise en service en 2003, compte trois fois plus de pixels (300 millions) et couvre la totalité du champ de vue utilisable au foyer primaire du CFHT, soit (0.92 deg^2) [Boulade et al. (2003)]. Son plan focal est composé de 40 CCD 2000×4000 de 40 microns d’épaisseur ($\epsilon(\lambda) > 0.7$ entre 4000 et 7000 Å, cf. Figure 3.2). À sa mise en service, MegaCam était la plus grande caméra astronomique CCD jamais construite. Elle a permis de réaliser un des grands relevés cosmologiques des années 2000-2010 : le CFHT-Legacy Survey.

Les générations suivantes ont permis d’une part d’accroître la taille des plans focaux, en préservant l’échantillonnage spatial, d’autre part d’améliorer significativement l’efficacité quantique dans le rouge grâce à la généralisation des CCD haute-résistivité. La caméra DECam (cf. Figure 3.4) développée pour le Dark Energy Survey (DES) comporte $5,85 \times 10^8$ pixels pour un champ de vue de 3 deg^2 [DePoy et al. (2008)]. Elle a été installée en 2013 sur le télescope Blanco (cf. Figure 3.13a) à l’observatoire interaméricain du Cerro Tololo (CTIO pour *Cerro Tololo Inter-American*

Observatory). La même année, la caméra HSC (pour *HyperSuprimeCam*) a été installée sur le télescope Subaru de 8.2m à Hawaï (cf. Figure 3.13b). Son plan focal comprend plus de 100 CCD [Miyazaki et al. (2012)] pour un champ de vue de 1.8 degrés carrés (880 millions de pixels). L'efficacité quantique des CCD de HSC est représentée par la courbe en bleu sur la figure 3.2, ils sont optimisés pour observer dans le rouge.

Enfin, le plus grand plan focal jamais construit est celui de la caméra qui équipera le télescope de l'observatoire Vera Rubin [LSST DESC (2012)]. Il est constitué de plus de 200 CCD 4000×4000 et totalise plus de 3 milliards de pixels. L'imageur couvre un champ de 9.2 deg^2 en préservant un échantillonnage spatial de 0.2 secondes d'arc par pixel. Cette caméra devrait voir sa première lumière en 2023.

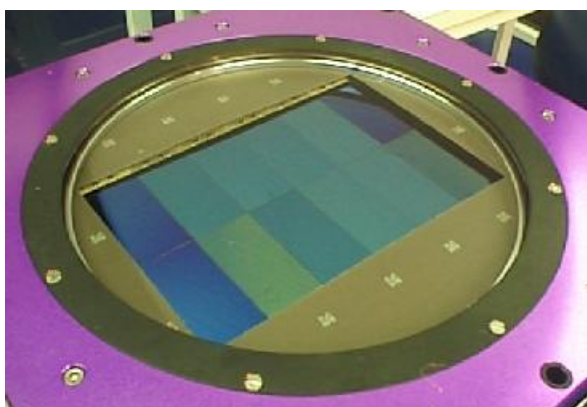


FIGURE 3.3 – plan focal de la caméra CFHT12K.



FIGURE 3.4 – plan focal de la caméra DECam.

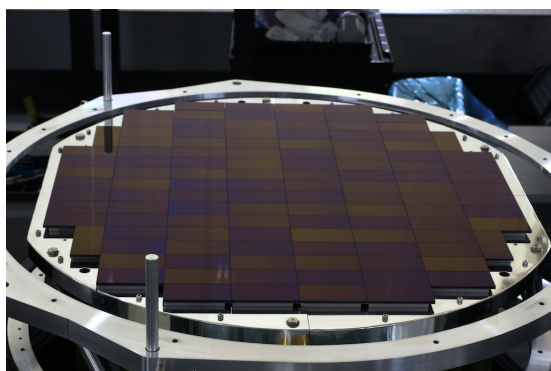


FIGURE 3.5 – plan focal de la caméra Hyper Suprime Cam (HSC)

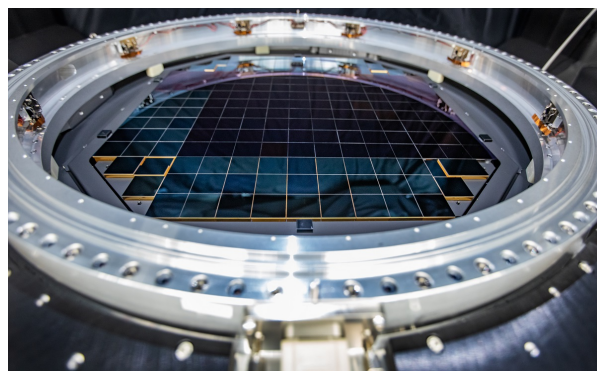


FIGURE 3.6 – plan focal de la caméra de l'observatoire Vera Rubin.

Les caméras grand champ présentées ci-dessus, permettent de collecter des images de grands volumes d'Univers, avec une excellente résolution spatiale (limitée par la qualité du site astronomique, généralement de l'ordre de la seconde d'arc). En revanche, l'information spectrale est minimale : chaque pixel contient l'intégrale du spectre de la lumière qu'il a collectée, dans une bande passante large (100–nm). Pour comprendre la physique des objets étudiés, ou pour mesurer leur décalage spectral, il est essentiel d'accéder à une information spectrale plus fine. C'est l'objet des observations spectroscopiques.

3.1.3 Observations spectroscopiques

Comparée à la photométrie, la quantité de photons nécessaire à l'acquisition d'un spectre est bien plus importante et nécessite un temps d'acquisition plus élevé (et/ou un miroir plus grand). Tous les instruments spectroscopiques en service actuellement utilisent des CCD comme senseurs. L'accroissement de la taille et de l'efficacité des CCD a permis de proposer des designs de spectrographes de plus en plus performants, permettant de collecter de plus en plus de spectres en parallèle. Dans cette section, nous présentons les principaux types d'instruments utilisés pour le suivi des SNe Ia.

Spectroscopie à fente (LSS pour *Long Slit Spectroscopy*).

Le type de spectrographe le plus répandu est le spectrographe à fente. Une fente est placée devant l'objet cible, afin de ne sélectionner qu'une direction spatiale. À ce stade, le faisceau est convergent. Une première optique permet de repasser dans le plan pupille. Le faisceau passe par un disperseur, généralement un grisme, puis est refocalisé sur un CCD. Le signal enregistré sur le CCD comporte une dimension spatiale (coordonnées le long de la fente) et une dimension spectrale (dispersion de la lumière orthogonalement à la fente). On peut voir sur la figure 3.7, à gauche la galaxie 19843 du suivi 3D-HST et à droite le spectre obtenu par LSS, avec une fente placée en $Y=0$. La dimension Y de pixel est remplacée par la longueur d'onde de la lumière. Cette technique a l'avantage d'être simple d'utilisation. En revanche, elle ne permet d'obtenir qu'un faible nombre de spectres par nuit, un seul objet étant observé à la fois. Il est possible d'augmenter le nombre de spectres obtenus par pose, grâce aux techniques de spectroscopie multi-objets (MOS pour *Multi Object Spectroscopy*). Un spectrographe MOS dispose de plusieurs fentes mobiles et orientables permettant de sélectionner plusieurs objets dans le champ de vue de l'instrument. La lumière est acheminée vers le système de dispersion soit par un système optique, soit par un faisceau de fibres optiques. Les spectres obtenus sont généralement enregistrés sur le même CCD.

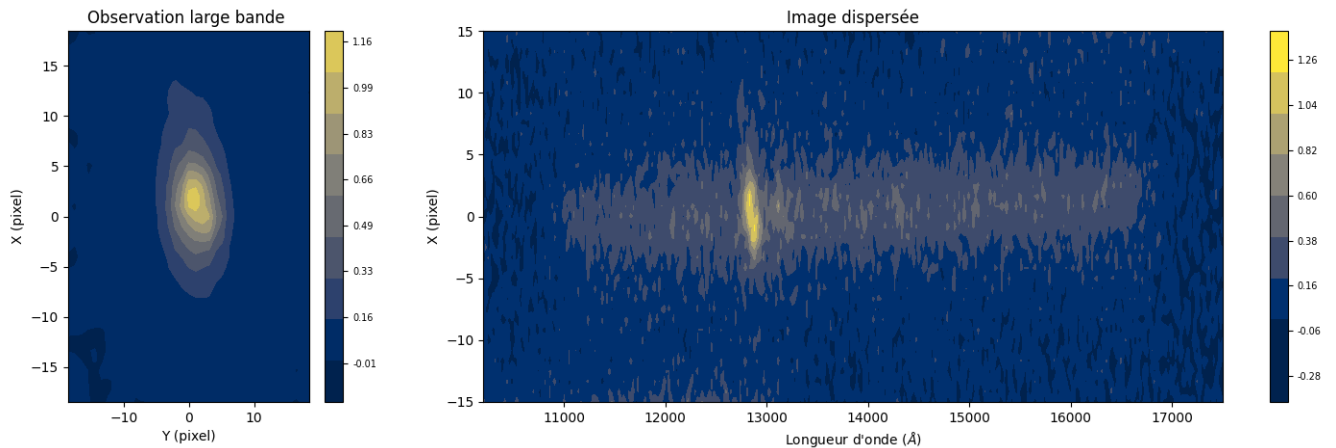


FIGURE 3.7 – À gauche, observation photométrique de la galaxie 19843 du suivi 3D-HST [Brammer et al. (2012)] et à droite l'image dispersée de cette même galaxie.

L'évolution technologique des spectrographes a suivi celle des caméras CCD. Le nombre de spectres et les procédés d'obtention ont beaucoup évolués sur les dernières années. Au début des années 2000, les spectrographes étaient des MOS à fente. Par exemple, l'instrument FORS1 était une caméra CCD, montée au VLT (*Very Large Telescope*) en 2005. Par un système de masques, 19 fentes pouvaient être placées simultanément sur le champ, dispersant la lumière comprise

entre 4000-11000 Å sur le plan focal. Les spectrographes MOS modernes peuvent collecter simultanément des milliers de spectres, grâce à des systèmes automatisés de positionneurs de fibres optiques. DESI [DESI Collaboration et al. (2016)], qui a vu sa première lumière en 2021, peut obtenir simultanément 5000 spectres en une observation. En 2023, le spectrographe 4MOST [de Jong et al. (2014)] pourra obtenir environ 2400 spectres simultanément.

Spectroscopie à intégrale de champ (IFU pour *Integral Field Unit*) – est une méthode qui permet d’obtenir le spectre de tous les pixels du plan focal de l’instrument. Le résultat est un cube de données (2 dimensions spatiales et une dimension spectrale). Il existe différentes façons de diviser le champ d’observation, dont trois principales sont représentées sur la Figure 3.8.

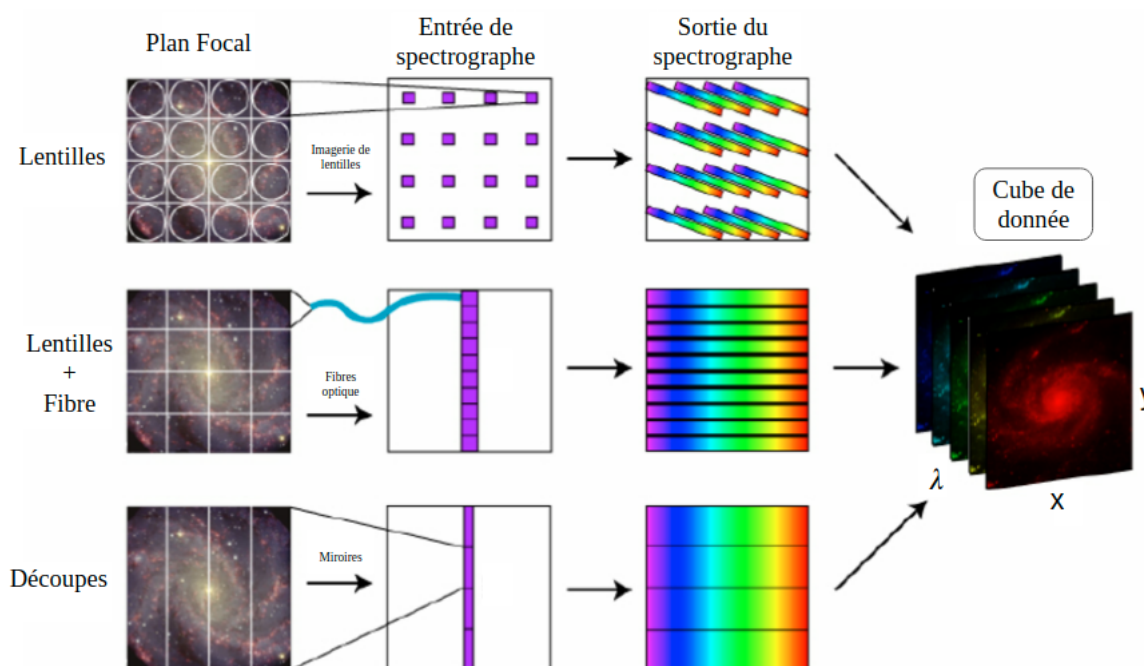


FIGURE 3.8 – Trois techniques de spectroscopie à intégrale de champ, figure adaptée de Westmoquette et al. (2009)

Une première méthode consiste à placer au plan focal une matrice de microlentilles. La lumière collectée par chaque microlentille est dispersée et les spectres refocalisés sont enregistrés sur un CCD. Une seconde méthode utilise des lentilles suivies de fibres optiques acheminant la lumière jusqu’à l’élément disperseur. Enfin, la dernière méthode utilise des miroirs pour découper le champ de vue en tranches (pseudo-fentes) qui seront ensuite dispersées et réarrangées sur un même CCD.

Un spectrographe à intégral de champ utilisé dans un relevé de SNe Ia est l’instrument SNIFS (*SuperNova Integral Field Spectrograph* Lantz et al. (2004)). SNIFS a été installé en 2004 sur le télescope de 2.2m de l’université d’Hawaï au Mauna Kea. Il a été utilisé pour le suivi des Supernovæ du programme *Nearby Supernovae Factory* (SNfactory). Son champ de vue $6'' \times 6''$ est divisé en 225 "spaxels" grâce à un réseau de 15×15 micro-lentilles (cf. Figure 3.9). Pour chaque spaxel, deux spectres sont obtenus. Une partie bleue entre 3200 Å et 5300 Å et un spectre rouge entre 5000 Å et 10200 Å. À la différence de la spectroscopie à fente, la totalité de la lumière du plan focal est collectée. Il n’y a pas de "pertes de fentes", et la réponse photométrique de l’instrument est donc contrôlable. La technique permet l’obtention de spectres dont la calibration

en flux a été montrée comme étant de l'ordre de quelque pour mille sur tout un réseau d'étoiles standard Rubin et al. (2022). On parle de spectro-photométrie. SNIFS a permis d'obtenir des données de suivi spectro-photométrique de grande qualité de plusieurs centaines de supernovæ détectées par la collaboration SNfactory (cf. Figure 2.6a).

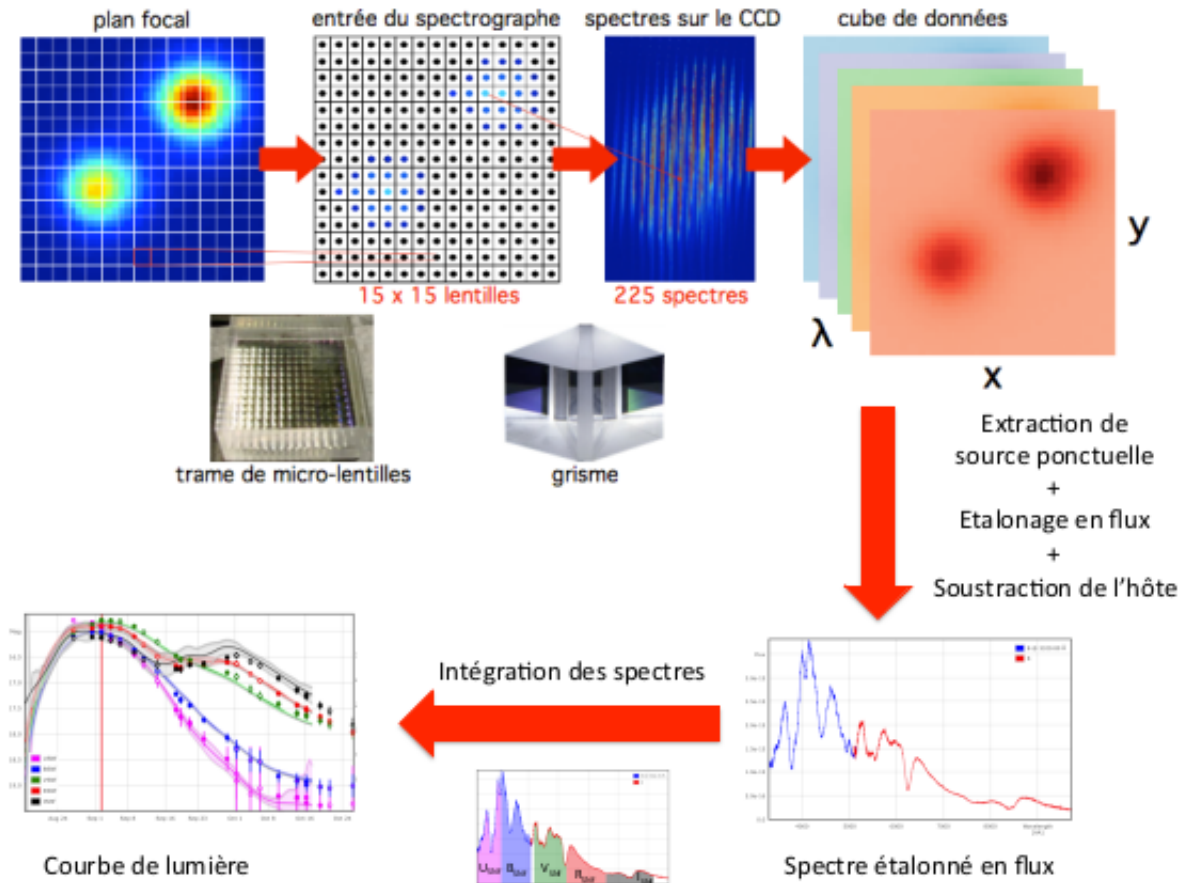


FIGURE 3.9 – Principe de fonctionnement de la spectroscopie à intégrale de champ réalisée par l'instrument SNIFS. Le plan focal est divisé en 225 spaxels. Pour chaque spaxel, deux spectres sont obtenus. De chaque champ, un spectre est obtenu permettant de constituer un cube de données à deux dimensions spatiales et une spectrale pour chaque pose. Les spectres des SNe en sont extraits. Le suivi réalisé permet de synthétiser des courbes de lumière pour ces objets. Figure extraite de Chotard (2011).

3.2 Données de suivi

Dans cette section, nous présentons les données de suivi qui permettent d'identifier les supernovæ de type Ia et de déterminer leur distance de luminosité. Deux sortes de données sont à notre disposition : des données photométriques, d'où on peut extraire le flux des SNe, intégré dans une bande passante large (O(100-nm)), puis des distances de luminosité et des données spectroscopiques, qui servent principalement au typage des détections et à l'obtention de leur *redshift*.

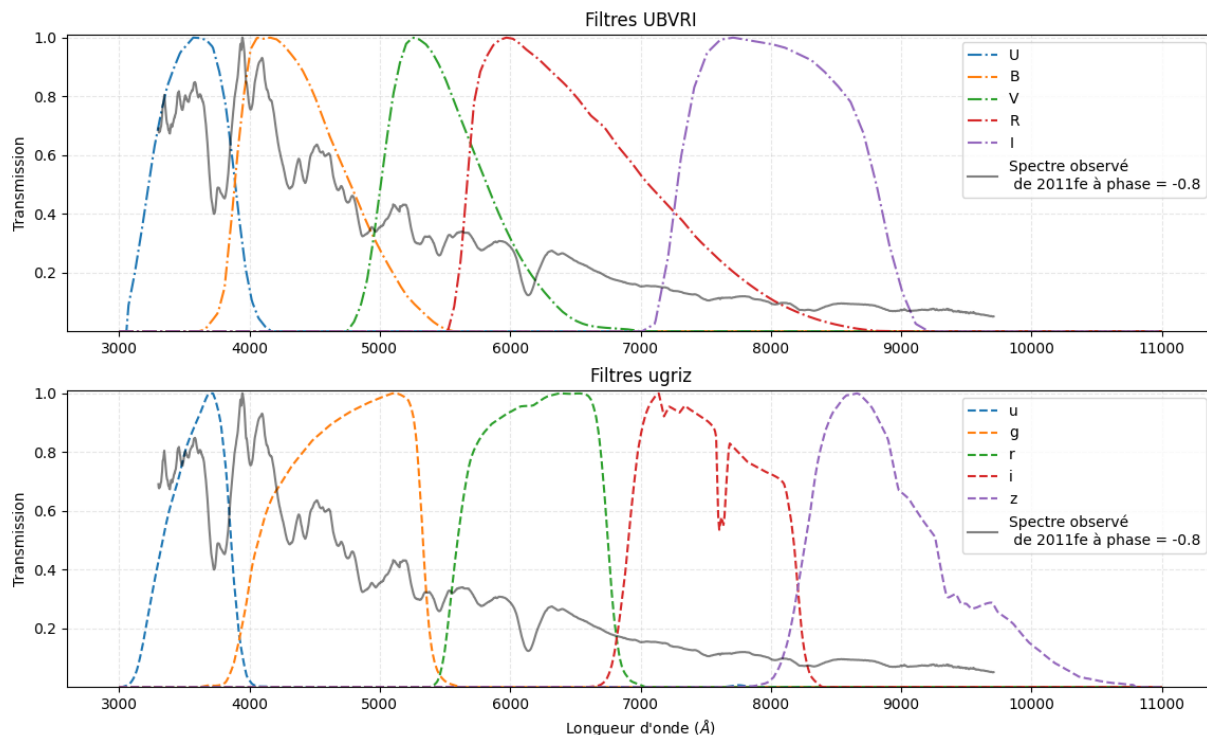


FIGURE 3.10 – Transmission des filtres *UBVRi* et *ugriz*, avec un spectre d’une SN observé à bas z proche du maximum de luminosité en bande B.

3.2.1 Photométrie en bande large

La photométrie est l’ensemble des techniques qui permettent de mesurer le flux d’une source ponctuelle ou étendue. Aujourd’hui, les instruments de photométrie sont presque tous des imageurs équipés de CCD. Afin de diminuer les fluctuations du fond de ciel, et d’obtenir une information partielle en longueur d’onde, on place généralement un filtre dans le chemin optique. Historiquement deux principaux systèmes de filtres ont été utilisés. Le système historique *UBVRi*, normalisé par Bessell (1990), a été utilisé par plusieurs générations d’instruments. Avec l’avènement des *coatings* interférentiels, permettant de réaliser des fonctions de transmission plus raides, un second type de filtres s’est imposé. Le système *ugriz*, défini pour le relevé SDSS [Fukugita et al. (1996)] a été adopté par tous les grands relevés cosmologiques (CFHT-LS, DES, Subaru/SSP, LSST...). On représente sur la figure 3.10, la transmission des filtres *UBVRi* et *ugriz*. Les premiers, réalisés en verre coloré, ont des coupures mal définies et les bandes passantes se recouvrent partiellement, à l’inverse des seconds.

En arrière plan, on peut voir le spectre de SN2011fe observé à une phase proche du maximum. Une mesure photométrique est l’intégrale du spectre $S(\lambda)$, multipliée par la transmission effective de l’instrument $T(\lambda)$. En général, S est exprimé en énergie par unité de temps, de surface et de longueur d’onde ($\text{erg/s/cm}^2/\text{nm}$). Les CCD comptent des photo-électrons. Le modèle de la mesure est donc :

$$\phi = A \int S(\lambda) \frac{\lambda}{hc} T(\lambda) d\lambda \quad (3.2)$$

où ϕ est le flux enregistré sur le CCD (en photo-électrons) et A est la surface du miroir. La transmission effective de l'instrument est le produit de plusieurs composantes : la transmission du filtre $T_{filt}(\lambda)$, la transmission de l'optique (réflectivité des miroirs et transmission des lentilles et fenêtres) T_{opt} , la transmission de l'atmosphère au moment de l'observation $T_{atm}(\lambda)$ et l'efficacité quantique du senseur $\eta(\lambda)$:

$$T(\lambda) = T_{filt}(\lambda) T_{opt} T_{atm} \eta(\lambda) \quad (3.3)$$

3.2.2 Analyse des données photométriques

La figure 3.11 présente une image du champ COSMOS produite par la caméra HSC. On distingue principalement des galaxies. Ce sont des sources étendues de formes complexes. On distingue également quelques étoiles, qui sont des sources ponctuelles. La lumière provenant d'une étoile est originellement une onde plane quasi-parfaite, distordue par les gradients d'indices $n(\lambda, t)$ générés par la turbulence atmosphérique, puis par les effets de diffraction à l'intérieur du télescope. Le flux d'une étoile est donc étalé sur plusieurs pixels. La fonction d'étalement est appelée *Point Spread Function* ou PSF. Sa forme est en bonne approximation, celle d'une gaussienne à deux dimensions. La largeur à mi-hauteur de cette gaussienne quantifie la qualité d'une image et s'appelle le *seeing*. La PSF est lentement variable sur le plan focal mais peut être fortement variable d'une image à l'autre. Au premier ordre, toutes les étoiles présentes sur une même image ont la même forme (modulo les variations spatiales de la PSF). Les supernovæ sont des sources ponctuelles. Dans la suite de ce chapitre, nous nous intéressons à la mesure du flux de ce type d'objets.

Les images brutes produites par le système d'acquisition du télescope contiennent des valeurs R_{ij} encodées en unités d'ADC ou ADU, qui est au premier ordre une fonction affine du flux de photons ϕ_{ij} en provenance du fond de ciel et des objets astrophysiques que l'on cherche à mesurer :

$$R_{ij} = \frac{1}{g} \times [T_{exp} t_{ij} \phi_{ij} + n_{ij}] + p_{ij} \quad (3.4)$$

g est le gain de la chaîne de lecture (en électrons par ADU), T_{exp} est le temps de pose et t_{ij} est la transmission du télescope pour le pixel ij . Le produit $T_{exp} t_{ij} \phi_{ij}$ est le nombre de photoélectrons collectés dans le pixel ij au cours de la pose. Ce nombre fluctue de pose à pose en suivant une statistique de Poisson. Par ailleurs, la mesure est affectée par un bruit généré par la chaîne de lecture et d'amplification. On modélise ces deux contributions par un bruit noté n_{ij} . Enfin, la chaîne de lecture de chaque canal est réglée de façon à ce qu'un signal nul corresponde à un niveau d'ADC suffisamment élevé pour pouvoir échantillonner les fluctuations dues au bruit de lecture. Cet *offset*, appelé piédestal est susceptible de varier avec le temps et doit être mesuré.

Les premiers éléments de la chaîne d'analyse consistent à extraire des images brutes, un signal qui soit une fonction linéaire du flux de photons original. Les piédestaux sont mesurés, généralement pour chaque ligne, via des mesures complémentaires réalisées par la chaîne de lecture après que chaque ligne du CCD a été enregistrée. Ces mesures complémentaires sont stockées dans les images brutes et sont appelées *overscan* pixels. Un modèle de piédestal p_{ij} est construit à partir des *overscans* et de séries d'images à temps d'exposition nul (*bias frames*) et soustrait à chaque image. Les variations spatiales de la réponse de la caméra sont mesurées à l'aide d'images d'illumination uniforme réalisées soit en observant le ciel en début ou fin de nuit (juste après/avant le coucher/lever du soleil), ou à l'aide d'un dispositif générant une illumination



FIGURE 3.11 – Observation photométrique du champs COSMOS, par l’instrument HyperSuprimeCam (HSC, de plan focale visible figure 3.5) du télescope Subaru (cf. Figure 3.13b)

uniforme. Ces images de champ plat, corrigées de leur piédestal, sont combinées pour former une image de correction F_{ij} , à haut signal sur bruit et normalisée à une valeur proche de 1 pour préserver la dynamique originelle du signal :

$$F_{ij} \propto \frac{1}{g} t_{ij} + n_{ij}^F \quad (3.5)$$

avec $\sigma_{n^F} \ll 1$. Ces deux corrections permettent de reconstruire, pour chaque image, un signal proportionnel au flux de photons incidents :

$$f_{ij} = \frac{R_{ij} - p_{ij}}{F_{ij}} \approx T_{exp} \phi_{ij} + n'_{ij} \quad (3.6)$$

Flux des sources ponctuelles Pour une source ponctuelle à la position (x_0, y_0) sur le CCD, le flux dans le pixel ij est donné par :

$$f_{ij} = f \times \psi(i - x_0, j - y_0) + b_{ij} + n_{ij} \quad (3.7)$$

où f est le flux de la source (la quantité que l’on cherche à reconstruire), $\psi(i - x_0, j - y_0)$ est la valeur de la PSF pour le pixel ij (en toute rigueur, l’intégrale de la PSF dans le pixel). b_{ij} est le flux du fond de ciel, moyenné sur la totalité de l’image et n_{ij} le bruit affectant la mesure de flux, dû aux fluctuations de Poissons du signal et du fond de ciel, ainsi qu’au bruit de lecture.

Estimateurs de Flux Plusieurs techniques d'estimation du flux des sources ponctuelles ont été développées. La plus simple et, en pratique, la plus robuste, consiste à sommer la valeur des pixels dans une ouverture de rayon r autour de l'étoile. Un tel estimateur permet d'obtenir une fraction du flux des objets. Connaître la valeur exacte de cette fraction importe peu, car elle est mesurée à l'étape de calibration. En revanche, il est important qu'elle soit : une fonction linéaire du flux, uniforme sur l'image et si possible, constante d'image à image. Afin de rendre l'estimateur aussi uniforme que possible, l'ouverture est souvent une fonction de la qualité d'image. En combinant les mesures dans plusieurs ouvertures différentes, il est possible de contrôler l'uniformité de l'estimateur avec des précisions de l'ordre du pour-mille. La photométrie d'ouverture n'est pas un estimateur statistiquement optimal. En revanche, elle ne requiert pas de ressources élevées en temps de calcul ni de modélisation fine de la PSF.

Lorsque le champ est encombré (forte densité d'étoiles) ou que la source ponctuelle que l'on cherche à mesurer est superposée à un fond structuré (supernova dans sa galaxie hôte) ou encore que la faiblesse de la source requiert un estimateur statistiquement efficace, on a généralement recours à un algorithme de photométrie de PSF. Pour un télescope au sol, la forme de la PSF n'est pas connue a priori et doit être modélisée à partir des étoiles de champ. Plusieurs approches ont été proposées pour la modélisation de la PSF. Une technique standard [Holtzman et al. (2008); Astier et al. (2013)] consiste à ajuster simultanément sur tous les objets du champ identifiés comme étoiles un modèle du type :

$$f_s \times [\psi(i - x_s, j - y_s) + r(i - x_s, j - y_s)] \quad (3.8)$$

où f_s , (x_s, y_s) sont les flux et la position de l'étoile s et où ψ est un modèle de PSF commun à toutes les étoiles. ψ est constitué d'une partie analytique, décrivant le coeur de la PSF et d'une partie empirique (grille de pixels) pour rattraper ce que la fonction analytique ne peut décrire. En général, ce modèle est lentement variable en fonction de la position sur le CCD.

L'estimation du flux des supernovae Les supernovæ apparaissent sur les images de suivi photométrique comme des sources ponctuelles faibles, superposées au fond, structuré, de leur galaxie hôte (cf. Figure 3.12). L'estimateur de flux doit répondre à deux prérequis, d'abord il doit pouvoir séparer le flux de la SN du flux sous-jacent de la galaxie et ensuite il doit être statistiquement efficace, en raison de la quantité limitée d'information présente dans l'image.

L'état de l'art dans le domaine est la famille d'algorithmes dits de *scene modeling* présentée dans Astier et al. (2013), qui consiste à ajuster par la méthode des moindres carrés, sur l'ensemble des pixels des images de suivi, un modèle décrivant la "scène" SN + galaxie (i.e. une source ponctuelle, superposée à une source étendue) :

$$\phi_{i,p} = [f_i \times \psi_i(x_p, T_i[x_{obj}]) + G(T_i^{-1}(x_p)) \otimes K_i + S_i] R_i \quad (3.9)$$

où $\phi_{i,p}$ est la valeur du pixel p de l'image i , ψ_i modélise la PSF sur l'image i , G est un modèle du fond de la galaxie et K_i est le noyau de convolution qui lie la PSF sur l'image i à la PSF sur l'image de référence (*seeing*) (image dite "de référence") : $\psi_i = \psi_r \otimes K_i$. x_p est la position de la SN sur l'image de référence. T_i est la transformation astrométrique entre l'image de référence et l'image courante. R_i est un coefficient photométrique liant les flux des sources ponctuelles sur l'image aux flux des mêmes sources sur l'image de référence. Les PSF, les noyaux de convolution et les transformations astrométriques et photométriques sont déterminées a priori, à partir des

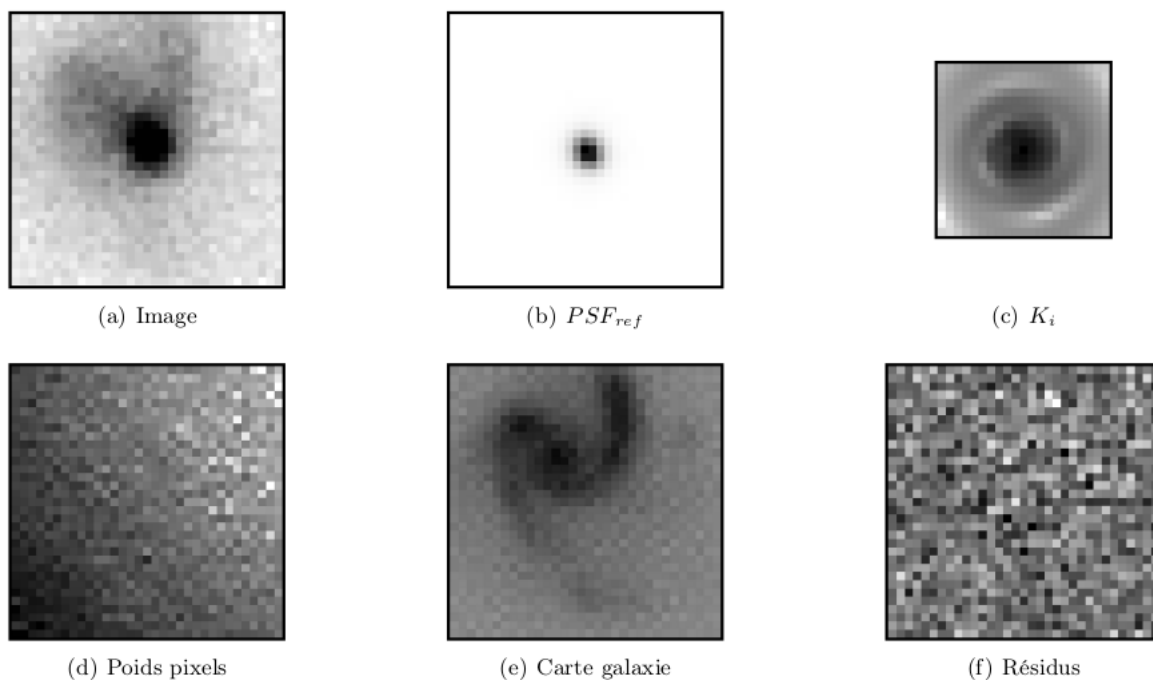


FIGURE 3.12 – Composants de l’ajustement du modèle de PSF pour une image dans la bande g_M proche du maximum de la SN Ia 03D4ag ($z = 0.285$) du relevé SNLS. Figure extraite de Fourmanoit (2010).

informations présentes dans les images de suivi. Le modèle est ajusté sur toutes les images disponibles dans une bande donnée et contenant la galaxie hôte (en pratique, sur un ensemble de vignettes centrées sur la position de la SN). Les paramètres libres du modèle sont les flux de la SN f_i , la position de la SN x_p et les paramètres du modèle de galaxie. Cela représente en pratique quelques milliers de paramètres (essentiellement les paramètres de la fonction G). Les seuls paramètres d’intérêt sont les flux f_i (quelques dizaines de paramètres). La figure 3.12 illustre l’ajustement du modèle sur une vignette de suivi.

Calibration des données photométriques Les algorithmes de photométrie décrits plus haut produisent des valeurs de flux exprimées en unités instrumentales (ADU) non calibrées. La relation entre le nombre de photo-électrons enregistrés au cours de la pose et le flux de photons au-dessus de l’atmosphère est une fonction de l’état du télescope et des conditions d’observation (réflectivité des miroirs, état des *coatings*, transparence de l’atmosphère).

La calibration a pour but d’exprimer les flux observés des supernovæ en flux physique, permettant ainsi de construire un diagramme de Hubble avec des flux de Supernovæ mesurés avec différents instruments. L’échelle absolue n’a pas besoin d’être connue car nous comparons des rapport de flux, il faut seulement déterminer l’échelle relative entre les différentes bandes d’observations. La magnitude d’un objet est une fonction du rapport du flux ϕ d’un objet observé à travers un instrument de transmission T et du flux d’un objet de référence de flux connu ϕ_{ref} :

$$m = -2.5 \log \left(\frac{\int \phi(\lambda) T(\lambda) \lambda d\lambda}{\int \phi^{ref}(\lambda) T(\lambda) \lambda d\lambda} \right) \quad (3.10)$$

$$m = -2.5 \log \left(\int \phi(\lambda) T(\lambda) \lambda d\lambda \right) + 2.5 \log \left(\int \phi^{ref}(\lambda) T(\lambda) \lambda d\lambda \right) \quad (3.11)$$

$$m = m^{inst} + ZP \quad (3.12)$$

m^{inst} est par définition la magnitude du flux observé par l'instrument. Le terme ZP s'appelle le point zéro, il définit l'échelle relative de flux. La calibration a pour but de déterminer sa valeur.

On utilise le flux d'étoiles connues, étoiles dites de référence, ou calibrateurs primaires. Différentes étoiles de référence peuvent être utilisées. Historiquement, Véga (α -Lyrae), dont les magnitudes dans les filtres de Bessel sont presque nulles, a été longtemps considérée comme étoile de référence. Cette étoile est très brillante² et par conséquent n'est pas observable par des instruments de nouvelle génération. Un autre calibrateur est l'étoile BD +17 4708, dont le spectre est connu à quelques pourcent près pour des longueurs d'onde allant de 1700 à 10000 Å. Si l'étoile de référence est observée dans un autre système de filtres, il faut préalablement déterminer les transformations pour passer d'un système de magnitude à l'autre, en observant des calibrateurs dit secondaires. Le projet CALSPEC [Bohlin (2014)] utilise le télescope Hubble pour mesurer le rapport entre un catalogue d'étoiles, dit CALSPEC, et des calibrateurs primaires. On observe ces étoiles isolées et brillantes présentes dans les champs d'observation de l'instrument.

En pratique dans une image, ni l'étoile de référence, ni les calibrateurs secondaires, du catalogue CALSPEC, ne sont présents dans les champs d'observations de SNe : on utilise donc des calibrateurs tertiaires. On constitue ainsi un catalogue qui donne les magnitudes des étoiles tertiaires pour pouvoir les comparer aux flux des SNe. On applique la même procédure d'extraction de flux aux étoiles de champ, pour avoir les mêmes biais et constituer un catalogue de calibration. Finalement, à partir d'un flux de SN en ADU, ϕ_{SN}^{ADU} , grâce à la moyenne des flux physiques, en ADU, de toutes les étoiles tertiaires $\phi_{tertiaire}^{ADU}$, puis en le comparant au flux du calibrateur primaire (dans le système de Bessel, ou dans le système *ugriz*, en passant par des calibrateurs secondaires), on obtient le flux physique de notre SN, ϕ_{SN} :

$$\phi_{SN} = \frac{\phi_{SN}^{ADU}}{\phi_{tertiaire}^{ADU}} \frac{\phi_{tertiaire}^{CALSPEC}}{\phi_{primaire}^{CALSPEC}} \phi_{primaire} \quad (3.13)$$

3.2.3 Données spectroscopiques

Considérons l'obtention d'un spectre par LSS, comme sur la figure 3.7. La valeur du spectre, sur le pixel de la ligne i et de la colonne j , vaut :

$$\phi_{ij} = \frac{1}{g} [f\psi_{i(\lambda)_j}(x - \bar{x}) + b_{ij}] * F_{ij} + p_i \quad (3.14)$$

Nous avons ici les mêmes éléments qu'en photométrie. Les étapes des corrections du piédestal et d'uniformisation du champ sont nécessaires. Il faut s'assurer que les longueurs d'onde de la

2. Vega est $10^6 - 10^8$ fois plus brillante que des SNe de *redshift* compris entre 0.3 et 1.

lumière dispersée soient calibrées. Pour cela, on illumine le plan focal à travers la fente avec, par exemple une lampe à arc, délivrant un spectre discret, dont les raies d'émissions, correspondant aux différents éléments présents dans le gaz, sont connues car mesurées en laboratoire. On peut ainsi associer à chaque pixel une longueur d'onde précise. La forme de la PSF dépend de la longueur d'onde. La lumière rouge sera moins sensible aux effets de perturbations atmosphériques que la lumière bleue. Ainsi, lorsque l'on place une fente ou une fibre devant la source, il est possible qu'une partie du flux ne soit pas prise en compte. En pratique, comme nous allons le voir, nous pouvons recalibrer les spectres, pour les corriger d'un éventuel effet chromatique. Enfin pour extraire le spectre de la SN de celui de la galaxie hôte, on ajuste par moindres carrés, la somme des spectres de référence d'un type de galaxie et d'un type de SN. Les paramètres de nuisance qui sont ajustés sont le *redshift*, la phase de la SN et la fraction de la galaxie contenue dans le spectre observé.

3.3 Les relevés

La constitution d'un échantillon de supernovæ de type Ia requiert de mettre en oeuvre des moyens de détection, de suivi photométrique et de caractérisation spectroscopique. Les supernovæ sont des événements rares : environ 10 000 explosent par an dans une sphère de rayon $z = 0.1$. La détection de plusieurs centaines de SNe dans la gamme de *redshift* $0 < z < 0.1$ requiert de pouvoir couvrir plusieurs milliers de degrés carrés avec une cadence de l'ordre de 4 à 5 jours au minimum, avec un télescope de 1 à 2 mètres de diamètre. La détection de quelques centaines de SNe plus distantes à $0.3 < z < 0.8$, nécessite de sonder plusieurs degrés carrés à une cadence analogue, avec un télescope de classe 4-m. Pour être efficace, le suivi spectroscopique requiert des télescopes de diamètre supérieur à ceux qui ont réalisé la détection et le suivi photométrique.

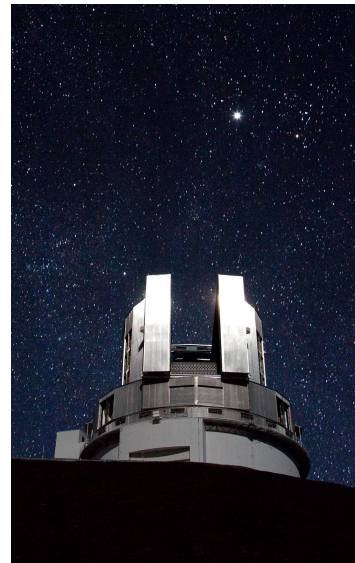
3.3.1 Première génération

Les premiers relevés de SNe Ia à pouvoir constituer des échantillons de quelques dizaines d'objets ont assemblé des moyens assez hétérogènes pour la détection et le suivi. L'instrument clef était un instrument grand-champ, utilisé pour la détection des candidats SNe. Le nombre de SNe détectées était fonction de la surface couverte et de la régularité de la cadence des observations de détection. Le typage spectroscopique et le suivi photométrique étaient réalisés par une multitude d'instruments plus petits. Cette génération regroupe les deux projets qui ont permis la découverte de l'accélération de l'expansion de l'Univers. Ils ont eu lieu entre les années 1990 et 2000. Un relevé proche mené par une équipe américano-chilienne a permis de découvrir quelques dizaines de SNe à des *redshifts* jusqu'à 0.18. Deux relevés réalisés avec des instruments de plus grand diamètre ont constitué indépendamment deux échantillons plus distants. C'est la combinaison des échantillons proches et distants qui a permis la détection de l'accélération de l'expansion de l'Univers.

- Le *Calán-Tololo Survey* [Hamuy et al. (1993b)] est un programme de suivi de supernovæ réalisé entre 1989 et 1995 à l'Observatoire inter-américain Cerro Tololo (CTIO) au Chili. Le projet a découvert et suivi 32 SNe Ia jusqu'à des *redshifts* de 0.1. La recherche était réalisée à l'aide de plaques photographiques, placées au foyer du télescope Blanco de 4-m (cf. Figure 3.13a). L'instrument avait un champ de vue de 15 deg^2 . Le suivi photométrique était ensuite réalisé avec une caméra CCD montée sur un télescope de 0.9m.
- Le *High-Z Supernova Search Team* [Riess et al. (1998)] est un programme qui a découvert et suivi 10 SNe Ia jusqu'à $z < 0.6$, en utilisant le télescope Blanco de 4-m situé au CTIO. Les observations photométriques ont été réalisées dans les bandes *B* et *V*. Les SNe ont été typées grâce à différents télescopes.



(a) Télescope Blanco de l'Observatoire inter-américain du Cerro Tololo (Chili)



(b) Télescope Subaru (Mauna Kea Hawaiï)



(c) Télescope Canada-France-Hawaiï (CFHT) (Mauna Kea Hawaiï)

- Le *Supernovae Cosmology Project* (SCP) [Perlmutter et al. (1999)] est un projet de suivi de supernovæ qui eut lieu à la même époque. 42 supernovæ de type Ia ont été découvertes à des *redshifts* allant de 0.18 à 0.83. Les SNe découvertes ont été suivies photométriquement à l'aide de cinq télescopes : le Blanco de 4m du CTIO, le 3.6m WYIN situé à l'Observatoire National de Kitt Peak aux États-Unis, le télescope de 3.6m de l'observatoire Européen austral (ESO) situé au Chili, le Isaac Newton Telescope (INT) de 2.5m et le télescope de 4.2m William Herschel, tous deux situés à l'Observatoire del Roque de los Muchachos aux îles canaries en Espagne.

Pour tous ces relevés, la stratégie de détection et de suivi, imposée par la faible disponibilité des instruments grand-champ, limitait la qualité de l'échantillonnage des courbes de lumière. Deux étapes de sélection reposant sur des décisions humaines étaient requises pour constituer la liste des candidats à un suivi photométrique, puis celle des candidats à un suivi spectroscopique.

Cette longue procédure, nécessitant l'utilisation de différents instruments, aboutissait souvent à des courbes de lumière amputées des données autour du pic de luminosité (notamment des don-

nées pré-maximum, région particulièrement intéressante). Le grand nombre d'instruments compliquait par ailleurs l'intercalibration des données photométriques. Enfin, la qualité de l'échantillonnage des courbes de lumière était limitée par le temps de télescope total disponible – souvent en quantité insuffisante. Par ailleurs, les lots de SNe produits par ces projets étaient affectés par des biais de sélection difficiles à contrôler. Pour certains, les SNe étaient recherchées à proximité de galaxies connues et préalablement sélectionnées. D'autre part, le processus de sélection des candidats à un suivi photométrique et spectroscopique, reposant quasi-intégralement sur les décisions des observateurs, était impossible à modéliser avec précision.

3.3.2 Génération 1.5

À la suite de la découverte de l'accélération de l'expansion de l'Univers, une série de projets de recherche et suivi de SNe s'est mise en place, dans le but d'augmenter significativement la statistique disponible. Certaines de ces initiatives ont conservé la même stratégie : instrument de détection distinct des instruments de suivi avec néanmoins nettement plus de temps de télescope disponible, ce qui a permis d'améliorer significativement la qualité des courbes de lumière. Les projets emblématiques de cette génération sont le *Carnegie Supernova Project* (CSP), un effort important mené par le *Harvard Center for Astrophysics* (CfA), ainsi que la *Nearby Supernova Factory*, qui a organisé son suivi autour d'un instrument innovant : le spectrographe intégral de champ SNIFS.

- Le *Carnegie Supernova Project* (CSP-I, Hamuy et al. (2006)) a découvert et suivi 134 supernovæ de type Ia [Krisciunas et al. (2017)] dont les *redshifts* sont compris entre 0.0037 et 0.0835. Les observations ont eu lieu entre 2004 et 2009. Ces SNe ont été observées par deux télescopes : Swope de 1m et duPont de 2.5m tous deux situés à l'observatoire de Las Campanas au Chili. Les observations photométriques ont été prises avec trois caméras CCD dans les filtres (*ugri* et *BV*) pour le visible et *YJHK* pour l'infrarouge. Les observations spectroscopiques ont été réalisées par LSS.
- Le *Center for Astrophysics* (CfA Hicken et al. (2007)) *Supernova survey* a observé, entre 2001 et 2011, 279 SNe dans la gamme de *redshift* $z < 0.08$. Les observations ont été réalisées avec le télescope de 1.2m de l'observatoire F. L. Whipple, avec 3 caméras³ pour la photométrie dans six bandes (*uUBVri*). Les observations spectroscopiques ont été réalisées par LSS avec le spectrographe FAST [Fabricant et al. (1998)]
- Le Berkeley SN Ia Program [Silverman et al. (2012)] a observé plus de 1200 spectres de 582 SN Ia de *redshift* inférieur à 0.2, entre 2009 et 2018. Les observations ont été réalisées pour l'essentiel avec le spectrographe Kast sur le télescope de 3m Shane à l'Observatoire Lick situé au Mont Hamilton aux États-Unis.
- Enfin, le programme appelé *Nearby Supernova Factory* (SNf Aldering et al. (2002)) a détecté et suivi, de 2004 à 2014, des SNe proches. Le suivi a été réalisé avec l'instrument SNIFS monté sur le télescope UH88-inch (de 2.2 mètres) de l'Université de Hawaï au Mauna Kea à Hawaï, permettant de couvrir chaque nuit 500 degrés carrés de ciel. L'utilisation d'un IFU a permis d'obtenir pour chaque supernova un suivi dit spectrophotométrique, i.e. dont la calibration en flux est maîtrisée au pourcent sur toute la gamme de longueur d'onde du spectrographe ($4000 < \lambda < 10000$). Plus de 300 supernovæ ont été découvertes et suivies par le projet, avec des *redshifts* compris entre 0.01 et 0.08, générant plus de 4000 spectres [Saunders et al. (2018)]. Le lot de données SNfactory est à ce jour le seul lot homogène comprenant plus de 300 séquences spectrales calibrées.

3. 4Shooter jusqu'en Aout 2004, puis Minicam jusqu'en juillet 2005 et ensuite KeplerCam

3.3.3 Seconde génération

Avec la disponibilité croissante des instruments grand champ, une stratégie innovante a été proposée, dite de recherche glissante, ou *Rolling Search*. Le principe est d'utiliser le même instrument grand champ pour la détection *et* le suivi photométrique. Chaque champ du relevé est observé à une cadence régulière, environ 3-5 jours, dans chaque bande. Chaque visite permet simultanément de découvrir de nouvelles supernovæ et de mesurer le flux des SNe déjà découvertes.

Cette stratégie s'est révélée particulièrement efficace, notamment dans la gamme de *redshift* $0.3 < z < 1$. Elle a permis de réaliser des observations de plusieurs SNe simultanément, d'où une économie importante de temps télescope : à un même instant, une dizaine de SNe dans la gamme $0.3 < z < 1$ sont en effet observables dans un champ de 1 degré carré. Une unique observation permet donc d'en détecter ou d'en observer plusieurs simultanément. La totalité des observations photométriques est désormais réalisée par un unique instrument, la seule contrainte étant la disponibilité de cet instrument pour fournir une cadence d'observation régulière. Une cadence *rolling search* permet également, en supprimant les contraintes de la détection ciblée et le délai entre la détection et le démarrage du suivi photométrique, d'améliorer la qualité des courbes de lumière, et ceci pour *toutes* les SNe découvertes. La principale amélioration constatée est l'échantillonnage avant le pic de luminosité.

Enfin, en supprimant l'étape de la sélection à un suivi photométrique, elle a supprimé une part importante de l'incertitude entachant la fonction de sélection du relevé. En revanche, la quantité de temps spectroscopique restant limitée, la sélection des candidats à une caractérisation spectroscopique reste un biais de sélection important, le principal dans l'incertitude sur la fonction de sélection de l'échantillon final.

Les relevés de cette génération organisés autour d'une stratégie de *rolling search* sont listés ci-dessous.

- Le SDSS-II [Frieman et al. (2008)] *Supernova Survey* a utilisé la caméra SDSS montée sur le télescope de 2.5m au Apache Point Observatory, au Nouveau Mexique, pour découvrir et suivre 532 SNe Ia avec des *redshifts* compris entre 0.013 et 0.47. Les observations se sont déroulées pendant 3 mois de septembre à novembre sur trois ans de 2005 à 2007. Le suivi a observé 300 deg² tous les deux jours dans 5 bandes d'observations (*ugriz*). Les spectres ont été obtenus grâce à différents instruments de différents télescopes⁴.
- Le relevé *SuperNova Legacy Survey* (SNLS, Astier et al. (2006)) a observé le ciel en utilisant la caméra MegaCam montée sur le CFHT (cf. Figure 3.13c), sur le Mauna Kea à Hawaï. Ce suivi a découvert 426 supernovæ à haut *redshift* ($0.082 < z < 1.06$). Les spectres ont été obtenus par le *Very Large Telescope* (VLT), quatre télescopes de 8 mètres situés dans le désert d'Atacama au Chili, représentant 240h d'observation de septembre 2003 à septembre 2005, les deux télescopes Gemini sud situé au Cerro Pachon et Gemini nord situé au Mauna Kea et enfin les télescopes Keck I et II, qui ont une surface totale de 10m, situés au Mauna Kea.
- Le relevé ESSENCE [Miknaitis et al. (2007)], a détecté et suivi 102 SNe dont les *redshifts* sont compris entre 0.1 et 0.78. Les observations photométriques ont duré quatre ans, de 2002 à 2006. Elles ont été réalisées avec la caméra MOSAIC montée au télescope Blanco (cf. Figure 3.13a) de 4m au CTIO dans deux bandes d'observation R et I. Les spectres ont été obtenus par une multitude de télescopes⁵.

4. A bas z :3.6m-NNT, 3.5m-ARC, 2.4m-MDM, 4.2m-WHT, 4m-KPNO, 2.6m-NOT et à haut z 9.2m-HET, 8.2m-Subaru, 10m-Keck et 11m-SAL.

5. Les télescopes : Keck I et II de 10-m, VLT de 8m Gemini nord et sud, le Magellan Baade de 6.5m, Tillinghast de 1.5m et les télescopes Clay et MMT.

- Le suivi Fondation [Foley et al. (2018)] a utilisé le premier télescope de Pan-Starr, appelé PS1, pour observer les trois quarts du ciel disponibles dans six bandes, $g, r, i, z, y, P1$. Ce télescope de 1.8m est situé au sommet du Haleakalā à Hawaï et a été mis en service le 6 décembre 2008. 800 SNe, découvertes par d'autres relevés, ont été suivies jusqu'à un *redshift* de 0.1.
- Le suivi *Pan-STARRS1 Medium Deep Survey* [Rest et al. (2014); Scolnic et al. (2018)] a découvert et suivi 279 SNe Ia dont les *redshifts* sont compris entre $0.03 < z < 0.68$. Les observations photométriques ont été réalisées avec le télescope PS1 et ses 6 filtres ($g, r, i, z, y, P1$). Quant aux observations spectroscopiques, elles ont été réalisées grâce aux programmes des télescopes MMT, Gemini, Magellan et Herschel.
- Le suivi du *Dark Energy Survey* (DES, Smith et al. (2020)) a observé 307 SNe de *redshift* allant de 0.017 à 0.85. Les observations se sont étalées sur 6 ans. L'instrument utilisé est la caméra DECam (cf. Figure 3.4) montée sur le télescope Blanco au CTIO avec quatre filtres *griz*. Les spectres ont été obtenus, pour la majorité⁶, par le spectrographe multifibres AAOmega/2dF composé de 392 fibres et installé sur le télescope de 3.9m anglo-australien (AAT).
- Enfin, le survey HSC-SSP (pour *Hyper Suprime Cam -Subaru Strategic Program*) est un suivi de SNe à très haut *redshift* réalisé par deux instruments. Le principal est la caméra HSC montée depuis 2004 sur le télescope de 8m Subaru au Mauna Kea à Hawaï (cf. Figure 3.13b). Sur deux saisons d'observation, 2016-2017 et 2019-2020, il a observé et suivi près de 500 SNe. Les spectres ont été obtenus par l'instrument FOCAS (pour *Faint Object Camera and Spectrograph*) monté sur le même télescope ainsi que les instruments FORS2 (VLT), AAOMEGA (AAT) et DESI (Mayall). La haute capacité de collection de photons dans le rouge et l'infrarouge de HSC (cf. Figure 3.2) en font l'instrument idéal pour observer des SNe à haut *redshift* jusqu'à environ $z \sim 1.1$. Au-delà, les photons émis dans les bandes B ne peuvent plus être observés avec un détecteur silicium. Pour étendre le diagramme de Hubble au-delà de $z \sim 1.1$, une stratégie a été mise en place : une sélection de SNe distantes ($1.1 < z < 1.4$) découvertes par le relevé a bénéficié de deux saisons d'observation dans l'infrarouge, réalisées avec l'instrument WFC3 du télescope spatial *Hubble*. Les données collectées par le projet sont en cours d'analyse.

3.3.4 Troisième génération

La stratégie de recherche glissante décrite ci-dessus, a permis d'augmenter considérablement la taille des lots de SNe cosmologiques, tout en améliorant l'échantillonnage et l'homogénéité de leur suivi photométrique. En revanche, tous ces échantillons sont affectés par un biais de sélection spectroscopique difficile à modéliser. Un seul projet à ce jour a été capable de résoudre ce problème, en se dotant de moyens d'identification dédiés.

- Le projet *Zwicky Transient Facility* (ZTF) en cours actuellement, a déjà détecté et suivi plus de 3000 SNe Ia [Dhawan et al. (2022)] en un peu plus de 2 ans d'opération. Les observations sont réalisées dans trois bandes *gri* par une caméra de très grand champ (47 deg^2) montée sur le télescope Ochin-Schmidt de 1.2m à l'observatoire Palomar. L'instrument est capable de couvrir de l'ordre de 3800 deg^2 par heure, soit l'intégralité du ciel extragalactique observable deux fois par nuit. Cela permet d'implémenter une recherche glissante à cadence élevée (1 à 3 jours) sur l'intégralité du ciel nord extragalactique. Le projet est doté d'un spectrographe intégral de champ dédié, monté sur le télescope de 60 cm du Palomar, appelé SEDmachine. Il est capable de typer spectroscopiquement l'intégralité des détections de magnitude $g < 18.5$. Un second spectrographe est actuellement

6. Les autres ont été obtenu avec les Gemini Nord et Sud, Keck, Magellan, MMT, SALT, VLT.

mis en service pour pousser cette limite à $g < 19$. Le lot produit par ZTF est en cours d'analyse. Il est complet spectroscopiquement jusqu'à un *redshift* de 0.06. Une première publication des données est prévue pour la fin de 2022.

3.3.5 Futurs relevés

Deux grands relevés de SNe sont prévus d'ici la fin de la décennie.

- Le futur grand relevé de supernovae de type Ia débutera en 2024. Il sera réalisé dans le cadre du *Large Survey of Space and Time* (LSST), à l'Observatoire Vera Rubin, actuellement en construction au nord du Chili. Le LSST observera pendant 10 ans, dans les six bandes *ugrizy*, l'intégralité du ciel austral, avec une cadence de 4 à 10 jours, suivant la bande [Ivezić et al. (2019)]. Le télescope en cours de construction a un diamètre de 8.4-m et est équipé d'une caméra de 9.6 deg^2 (cf figure 3.6).

Le projet LSST comprend deux composantes principales : un relevé dit "Wide, Fast, Deep" (WFD) qui représentera près de 90% du temps total du programme et consiste en des observations cadencées peu profondes (30 s par visite) de tout le ciel austral, et une série de champs spécifiques dits "Deep Drilling Fields" (DDF) beaucoup plus profonds (600 s par visite). Chacune de ces deux composantes a une cadence choisie pour permettre d'implémenter une rolling search dans la gamme de *redshift* $0.08 < z < 0.4$ pour le WFD et $0.4 < z < 1$ pour les DDF. En dix ans d'opérations, LSST est susceptible de produire $O(10^5)$ SNe dans le WFD et $O(10^4)$ SNe dans les DDF. Ces estimations, ainsi que la qualité des courbes de lumière dépendent fortement de la cadence d'observation qui sera choisie par le relevé [Gris et al. (2022)].

LSST ne disposera vraisemblablement pas d'accès à des ressources spectroscopiques suffisantes pour identifier systématiquement les SNe Ia, même dans une gamme de *redshift* limitée. Le relevé fonctionnera dans un mode proche de celui de DES ou de HSC/Subaru : les SNe Ia seront identifiées photométriquement, à partir de leurs courbes de lumière. Les *redshifts* des galaxies hôtes seront obtenus a posteriori au moyen de spectrographes multifibres comme 4MOST [de Jong et al. (2019)] composé de 2436 fibres, installé sur le télescope VISTA au Chili.

- Un autre futur relevé important de SNe Ia est celui qui sera réalisé par le télescope spatial *Nancy Grace Roman Space Telescope* [Rose et al. (2021)]. Son lancement est prévu pour mai 2027. Il sera équipé d'une caméra grand champ de 0.28 deg^2 , équipée de six filtres *R, Z, Y, J, H* et *F* couvrant un intervalle de longueur d'onde de 6200 à 20000 Å. La cadence décrite dans Hounsell et al. (2018) est de 6 mois d'observation étalés sur 2 ans, avec 30h d'observation par visite tous les cinq jours. Les observations seront divisées en 3 sous-relevés. Un premier relevé large de 27.5 deg^2 , observant des SNe jusqu'à $z \sim 0.4$ dans les bandes *Y* et *J*. Un second relevé plus restreint de 9 deg^2 observant dans les bandes *J* et *H* mais pouvant aller jusqu'à $z < 0.8$. Enfin un profond de 5 deg^2 observera dans les bandes *J* et *H*, des SNe à haut *redshift* < 1.7 . Le suivi devrait observer plus de 20000 SNe Ia sur toute la durée de l'expérience, dont 12 000 avec un bon échantillonnage temporel soit : une observation avant les premiers cinq jours à partir du maximum de luminosité, une observation après 15 jours et dont les flux estimés par simulations ont un signal sur bruit supérieur à 5 dans 2 bandes [Rose et al. (2021)].

Hormis LSST, aucun des lots de données listés plus haut ne couvre une gamme de *redshift* suffisante pour réaliser une mesure de w . Ils doivent être combinés. Betoule et al. (2014) a produit un diagramme de Hubble de 740 SNe, provenant des relevés SNLS, SDSS, CfA, CSP et quelques unes de HST. Le diagramme de Hubble publié dans Scolnic et al. (2018) comprend les

1048 SNe (cf. Figure 3.14) des échantillons CfA, Panthéon et DES.

	SNe (Δz)	LC	référence	Spectra	référence
SDSS	484 (0.01 - 0.47)	2405	Holtzman et al. (2008), Sako et al. (2014)	658	Sako et al. (2014)
SNLS	412 (0.08 - 1.00)	1630	Betoule et al. (2013)	427	Balland et al. (2009), Howell et al. (2005), Bronder et al. (2008), Ellis et al. (2008), Walker et al. (2011)
CSP	92 (0.005 - 0.084)	522	Krisciunas et al. (2017)	630	Folatelli et al. (2013)
CFA	192 (0.006 - 0.08)	868	Hicken et al. (2009), Hicken et al. (2012)	1455	Blondin et al. (2012), Matheson et al. (2008), Branch et al. (2003), Krisciunas et al. (2011), Sauer et al. (2008), Li et al. (2003), Wang et al. (2009), Hicken et al. (2007)

TABLE 3.1 – Données photométriques et spectroscopiques des SNe Ia (de *redshift* connu) observées et suivies par CFA, CSP, SNLS et SDSS, à notre disposition. Un travail de récolte, de compréhension et de reformatage de ces données a été réalisé.

3.4 Diagramme de Hubble - contraintes sur l'équation d'état de l'énergie noire

3.4.1 Mesure de distance avec des SNe Ia

L'inférence des paramètres cosmologiques consiste à ajuster sur les mesures des modules de distance des SNe Ia une prédiction théorique de la relation module de distance-luminosité, noté μ_{th} :

$$\mu_{th}(z, \Omega_m, \Omega_\Lambda, w_0, w_a) = m(z, \Omega_m, \Omega_\Lambda, w_0, w_a) - M_B = 5 \log \left(\frac{d_l(z, \Omega_m, \Omega_\Lambda, w_0, w_a)}{10 \text{pc}} \right) \quad (3.15)$$

Les modules de distance mesurés sont standardisés ce qui permet de réduire significativement (0.4 mag \rightarrow 0.15 mag) la dispersion dans le diagramme de Hubble. Plusieurs techniques de standardisation ont été proposées. La plus courante est la standardisation proposée par Tripp & Branch (1999) :

$$\mu = m_B^* - \tilde{M}_B + \alpha X_1 - \beta c + \Delta_M \quad (3.16)$$

où les paramètres α (brighter-slower), β (brighter-bluer) et \tilde{M}_B sont ajustés simultanément avec

la cosmologie. Les paramètres m_B^* , X_1 et c sont inférés par l'ajustement d'un modèle empirique du flux des SNe.

3.4.2 Modèle empirique d'ajustement de courbes de lumière

En pratique, une fois que nous avons des spectres et courbes de lumière pour les SNe à différents *redshifts*, il faut transformer cette information en distances standardisées. Il faut extraire pour toutes les SNe, la magnitude en bande B dans le référentiel de la SN, bande notée B^* , un paramètre de forme ainsi qu'une couleur dans le référentiel de la SN. Ni la magnitude en bande B^* , ni la couleur mesurée dans le référentiel de la SN ne sont directement mesurées. L'observateur n'a accès qu'à des flux (ou des magnitudes) qui mesurent l'intégrale du spectre redshifté dans les bandes passantes du télescope.

Nous devons donc disposer d'un modèle spectral qui permet d'inférer à partir des flux mesurés, le flux et la couleur exprimés dans le référentiel de la supernova. En pratique, on cherche à construire un modèle de l'évolution spectrophotométrique de la SN, qui puisse prédire la courbe de lumière mesurée de la SN en fonction des paramètres que l'on cherche à reconstruire. Les modèles physiques décrivant l'explosion des SNe ne donnent actuellement que des prédictions qualitatives. L'état de l'art consiste à construire à partir d'un ensemble de SNe bien mesurées, un modèle empirique de l'évolution spectrophotométrique des SNe Ia. L'élaboration d'un tel modèle constitue le coeur de ce travail de thèse. Nous donnons dans le chapitre 4 une revue rapide des modèles proposés au cours des deux dernières décennies. Les chapitres 5 - 9 décrivent notre travail sur le sujet.

3.4.3 Diagramme de Hubble

Une fois les magnitudes B^* et les paramètres de forme et de couleur extraits des courbes de lumière des supernovæ, l'inférence des paramètres cosmologiques suit un schéma assez classique. La procédure consiste en un ajustement par moindres carrés. On minimise le χ^2 suivant :

$$\Delta\chi^2 = \Delta\mu^T V^{-1} \Delta\mu \quad (3.17)$$

avec $\Delta\mu = \mu_{th}(z_i; \Omega_m, \Omega_\Lambda, w_0, w_a) - \mu_i(m_B, X_1, c; \alpha, \beta, \tilde{M}_B)$

où V est la matrice de covariance des données. Lors de cet ajustement les paramètres cosmologiques sont déterminés en même temps que les paramètres définissant les lois de standardisation $\alpha, \beta, \tilde{M}_B$ et Δ_M . La pondération est la somme des erreurs à propager lors de l'ajustement et sera décrite en détail dans la section suivante. La méthode originale SALT2 déterminait α, β dans le même ajustement que les paramètres cosmologiques. Mais de nombreux résultats récents (Pantheon, DES, Pantheon+) utilisent la méthode de [Marriner et al. \(2011\)](#) (au sein de de *BEAMS with Bias Corrections* (BBC) [[Kessler & Scolnic \(2017\)](#)]) pour déterminer α, β indépendamment de la cosmologie, et produire un diagramme de Hubble corrigé des biais. La figure 3.14 montre le diagramme de Hubble extrait de [Scolnic et al. \(2018\)](#), composé de 1048 SNe découvertes et suivies par 12 relevés différents. Les valeurs des paramètres cosmologiques extraites de cet ajustement du diagramme de Hubble sont :

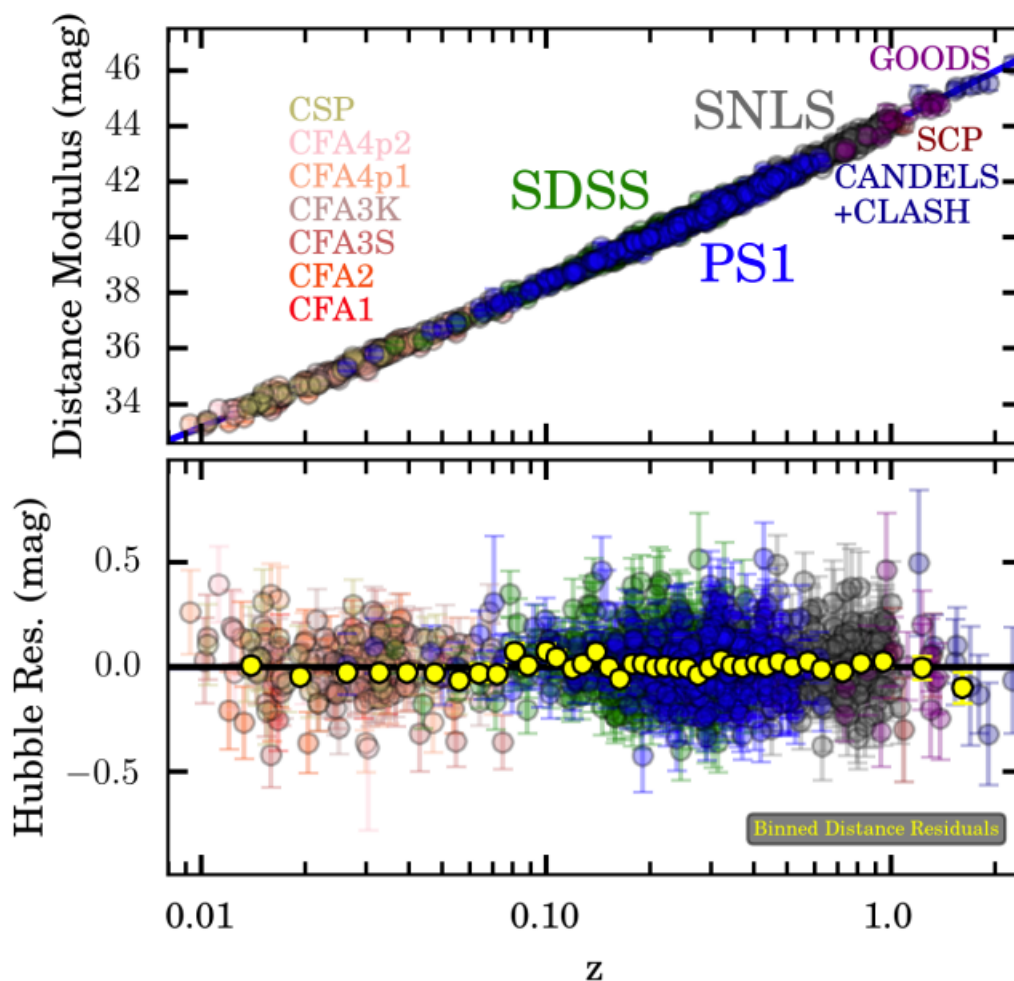


FIGURE 3.14 – Diagramme de Hubble extrait de l’analyse Panthéon Scolnic et al. (2018). Combinaison de 1048 SNe de *redshift* compris entre 0.03 et 2.3, observées par 12 relevés différents.

$$\Omega_m = 0.307 \pm 0.012 \text{ \& } w = -1.026 \pm 0.048 \quad (3.18)$$

Ces valeurs sont compatibles avec une constante cosmologique ($w = -1$). De nombreuses sources d’incertitudes statistiques et systématiques entachent cette mesure, qui doivent être propagées jusqu’au résultat final. La figure 3.15 détaille le bilan des incertitudes du résultat publié par Scolnic et al. (2018). Nous en discutons les principales contributions dans la section suivante.

3.5 Sources d’incertitudes

Le bilan des incertitudes affectant les mesures des paramètres cosmologiques peut classiquement se diviser en deux parties : les incertitudes statistiques, qui décroissent avec la quantité de données disponibles (nombre de SNe, nombre de mesures par SN, qualité photométrique des mesures) et les incertitudes systématiques qui affectent l’instrumentation, la stratégie de prise de données ou d’analyse, et ne décroissent pas en ajoutant des SNe ou en améliorant la qualité du suivi.

Les incertitudes statistiques et systématiques affectant les modules de distances sont décrites par leur matrice de covariance respective : V^{stat} et V^{sys} dont la somme compose la matrice V :

$$V = V^{stat} + V^{sys} \quad (3.19)$$

V^{stat} et V^{sys} sont elles même la somme de plusieurs contributions qui sont détaillées dans les sections qui suivent.

	w shift	σ_w^{syst}	Fraction of $\sigma_w^{(stat)}$
Stat. Uncertainty	+0.000	0.031	1.000
Total Sys Uncertainty	+0.031	0.025	0.814
Calibration			
SALT2 Cal	-0.002	0.014	0.457
Survey Cal	+0.006	0.009	0.285
HST Cal	-0.006	0.006	0.177
Supercal	+0.002	0.003	0.098
SN Modeling			
Selection	+0.010	0.007	0.233
Intrinsic Scatter	+0.019	0.005	0.170
β Evol.	-0.001	0.007	0.238
γ Evol.	-0.002	0.000	0.000
m_{step} Shift	-0.002	0.002	0.064
External			
MW Extinction	+0.010	0.008	0.262
Pec. Vel.	+0.000	0.003	0.103

FIGURE 3.15 – Proportion d'erreurs systématiques sur l'estimation de l'équation d'état de l'énergie noire [Scolnic et al. (2018)]. La source principale d'erreurs systématiques est la propagation des erreurs de calibration dans le "modèle d'ajustement de courbes de lumière" utilisé nommé SALT2 (dont on verra, le but, le développement, les réussites et les limitations dans la suite du manuscrit).

3.5.1 Incertitudes statistiques

Les incertitudes statistiques sont contenues dans la matrice V^{stat} , une matrice diagonale dont les valeurs sont le dénominateur de l'équation 3.17. Elles proviennent de plusieurs sources :

$$V^{stat} = V_N + V_z + V_{lens} + V_{int} \quad (3.20)$$

Incertainitudes photométriques Le terme V_N tient compte des bruits d'origine instrumentale provenant de la photométrie des objets. Certains termes dépendent du flux de l'objet. La mesure du flux de photons sur un CCD implique la présence d'un bruit de grenaille, modélisé par un processus de Poisson. A cela s'ajoute le bruit de lecture et d'amplification, indépendant du flux et négligeable pour des temps d'exposition long. Ces deux termes sont regroupés dans le bruit de variance σ_{RON}^2 . À ce bruit s'ajoute les contributions du fond de ciel et de la galaxie hôte qui sont regroupés dans le terme f_{sky} et la variance sur le flux de la SN f_{SN} , tout deux suivant des distributions de Poisson :

$$V_N(f) = \mathbb{E}[f_{SN}] + \mathbb{E}[f_{sky}] + \sigma_{RON}^2 \quad (3.21)$$

Pour les SNe Ia proches, les contributions du fond de ciel et de la galaxie hôte sont négligeables, le bruit varie comme la racine carrée du temps d'exposition. Concernant les SNe plus lointaines, la contribution du fond de ciel domine.

Vitesses particulières Le *redshift* mesuré est la somme de l'effet de l'expansion de l'Univers et d'une contribution propre des galaxies. Cet effet est d'autant plus grand pour les supernovæ proches et induit une incertitude de 0.1 mag pour une SN à $z_{lim} \sim 0.01$. Il est négligeable pour les SNe les plus lointaines. C'est pourquoi les supernovæ à z inférieur à cette limite sont exclues des analyses cosmologiques. Il est de plus possible de corriger des vitesses particulières pour des *redshifts* inférieurs à 0.06 en prenant compte le champ de vitesse calculé dans le groupe local, de distribution normale avec $c\sigma_z^{SN} = 150 \text{ km.s}^{-1}$ [Conley et al. (2011)]. Cette incertitude est propagée à l'incertitude du module de distance par la formule suivante, en notant σ_z , la racine carrée de la diagonale de la matrice V_z :

$$\sigma_z = \frac{5\sigma_z^{SN}}{z c \log(10)} \quad (3.22)$$

Lentillage gravitationnel La distribution de masse dans l'Univers est inhomogène. La lumière des SNe peut être affectée par des effets de lentille gravitationnelle du fait de la présence de halos de matière noire sur la ligne de visée, compliquant les tentatives de standardisation. Les photons émis par différentes SNe ne suivent pas le même trajet et sont affectés d'effets de lentillage gravitationnel dépendant de leur parcours.

L'intensité d'une supernova peut donc être augmentée, on parle de *magnification* ou au contraire diminuée c'est alors de la *dé-magnification*. Quelques SNe sont magnifiées mais la plupart sont faiblement démagnifiées. Cet effet augmente avec la distance parcourue, donc avec le *redshift*. L'incertitude sur la mesure du module de distance due à cet effet est donc une fonction du *redshift* et a été déterminée dans Jönsson et al. (2010) :

$$\sigma_{lens}(z) = 0.055 z \quad (3.23)$$

σ_{lens}^2 représente la diagonale de la matrice V_{lens} .

Dispersion intrinsèque de la magnitude des SNe La matrice V_{int} tient compte du fait que la standardisation α et β ne prend pas en compte la totalité de la variabilité des SNe Ia. Après standardisation, la dispersion des résidus au diagramme de Hubble est de l'ordre de 14%. Cette dispersion intrinsèque est la plus grande source d'erreur statistique. Les modèles de flux de supernovæ SNEMO [Saunders et al. (2018)] et SUGAR [Léget et al. (2020)], considérant plus de paramètres de standardisation, montrent une dispersion intrinsèque de la magnitude des SNe de l'ordre de 10%. D'autres méthodes ont montré des standardisations de 8% [Fakhouri et al. (2015)] et de 7% [Boone et al. (2021)]. Ces techniques nécessitant des observations de spectres au maximum de luminosité à haut signal sur bruit ne sont cependant pas directement applicables aux échantillons cosmologiques actuels.

3.5.2 Incertitudes systématiques

Les incertitudes systématiques proviennent de plusieurs sources :

$$V^{sys} = V_{biais} + V_{MW} + V_{cal} + V_{modele} + V_{hote} \quad (3.24)$$

Les proportions relatives de ces différentes sources aux incertitudes statistiques totales, extraites de l'analyse Panthéon sont résumées sur la figure 3.15. On distingue les incertitudes systématiques instrumentale et celle lié a la modélisation du flux des SNe.

Incertitudes systématiques liées aux techniques d'observation

Biais de sélection Le terme V_{biais} prend en compte le fait qu'un relevé de supernovæ est limité en flux et imposé par la sensibilité des instruments. Ainsi, il existe un biais de sélection favorisant l'observation des objets les plus lumineux, nommé biais de *Malmquist*. On peut voir sur la figure 3.16 que ce biais favorise la détection des SNe Ia les plus lumineuses ce qui correspond respectivement à des paramètres de couleur c négatifs (SNe Ia les plus bleues) et des paramètres d'évolution temporelle X_1 les plus grands (SNe Ia les plus lentes). Ainsi ce biais affecte directement la standardisation, il est particulièrement visible à haut *redshift*.

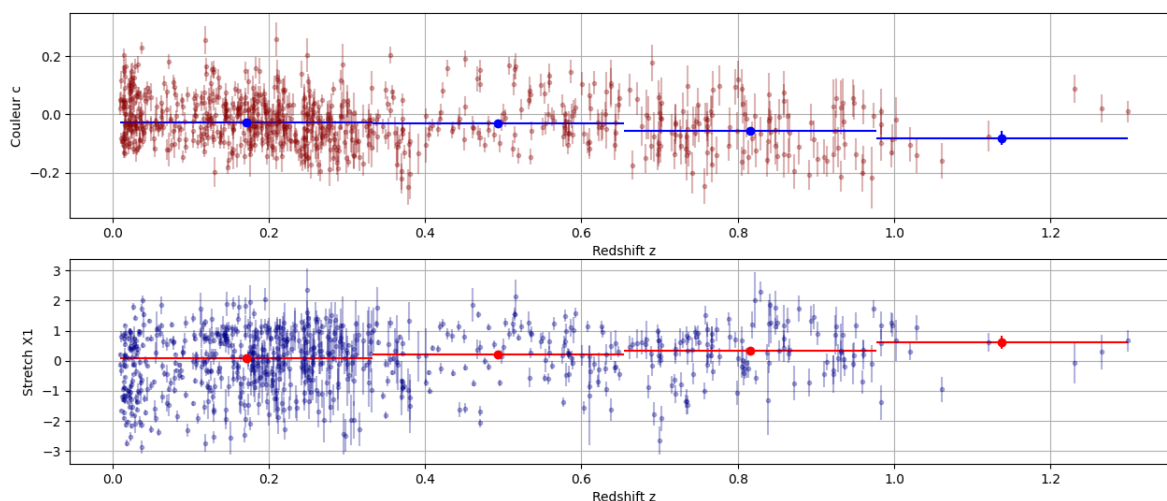


FIGURE 3.16 – Le biais de Malmquist affecte les distributions de stretch et de couleur des SNe à grand *redshift* de JLA.

Néanmoins puisque les SNe ont une dispersion en magnitude relativement faible, l'impact de ce biais est limité. La quantification de ce biais peut être estimée en fonction du relevé et du *redshift* grâce à des simulations. L'étude de la chaîne de détection des SNe peut être utilisée sur des données synthétiques, dont la distribution des différents paramètres est connu à priori et dont la comparaison avec les distributions reconstruites permet de quantifier l'introduction de biais par les différentes étapes. Il engendre une incertitude finale sur les résidus au diagramme de Hubble de l'ordre de 0.05 mag [Kessler et al. (2013)]

Extinction dans la Voie Lactée La lumière provenant des SNe Ia subit une extinction due à la poussière de la Voie Lactée dont il faut quantifier l'effet. La matrice V_{MW} tient compte de l'incertitude sur cette extinction de la lumière par la Voie Lactée. Des cartes d'extinction ont été définies en fonction de la ligne de visée [Cardelli et al. (1989), Schlegel et al. (1998) et Schlafly & Finkbeiner (2011)]. Cette extinction touche néanmoins davantage les SN à bas *redshift*, car observées dans le bleu, là où l'extinction est la plus forte. S'ajoute à cela le fait qu'à bas *redshift*, la nécessité de couvrir une large partie de ciel rend impossible de se cantonner aux zones de basse extinction galactique. En revanche pour les relevés à bas z , pour obtenir un volume suffisant il est nécessaire d'observer de grandes parties du ciel. Il est donc difficile d'éviter ces régions.

Calibration L'une des principales sources d'incertitude sur les mesures de distances de SNe est l'incertitude de calibration, V_{cal} . Elle provient des erreurs de mesure de flux des différents standards de calibration et de leur propagation selon la chaîne de calibration utilisée, permettant de comparer les SNe à bas et à haut *redshift*. Ainsi, cette erreur de calibration est la principale source d'erreur de JLA dans l'estimation des distances et donc des paramètres cosmologiques. Son amélioration est un des grands enjeux pour les relevés des prochaines expériences. La propagation des incertitudes de calibration au travers du modèle d'ajustement de courbe de lumière est la principale source d'incertitude systématique sur l'estimation de w (cf. Figure 3.15). Au total, les incertitudes de calibration de chaque relevé, de leur intercalibration et de leur propagation dans le modèle d'ajustement de courbes de lumière représentent plus de 56% de l'incertitude totale sur le paramètre d'équation d'état de l'énergie noire.

Modélisation spectrophotométrique Les différents modèles d'ajustement de courbes de lumière ont tous été entraînés sur des lots d'entraînement différents mais ayant pour point commun d'être de taille finie, d'être affectés par un bruit de mesure non négligeable, et enfin de ne pas couvrir l'ensemble de l'espace des paramètres. Le modèle est donc affecté d'une incertitude statistique. Comme il est utilisé pour ajuster l'ensemble des SNe du lot cosmologique, cela induit une incertitude systématique sur la détermination des paramètres d'intérêt des SNe.

À l'aide de simulations, Mosher et al. (2014) montre que pour des échantillons réalistes comme celui obtenu par JLA, les paramètres décrivant les SNe Ia sont biaisés. Un biais dans la description du flux des SNe Ia dans l'UV est de plus mis en évidence.

Incertitude de standardisation V_{hote} représente la contribution de l'incertitude sur la standardisation de SNe Ia. En effet, des signes de dépendance entre la luminosité observée et des propriétés de la galaxie hôte ont été observés. Il faut donc prendre en compte l'incertitude sur la paramétrisation de cette dépendance. Cette systématique correspond à plusieurs sources d'incertitude du tableau visible sur la figure 3.15 : la dispersion intrinsèque des SNe, la sélection et une possible partie de l'incertitude sur l'évolution de la loi de Brighther-Bluer, β , si une dépendance entre les propriétés de la galaxie hôte et la couleur sont observées.

3.6 Conclusion

Depuis une dizaine d'années, les incertitudes statistiques et systématiques affectent à parts égales les mesures de w issues du diagramme de Hubble des SNe Ia. Dans Betoule et al. (2014), les incertitudes statistiques représentent la moitié du bilan d'incertitude global. Dans l'étude publiée par Scolnic et al. (2018), elles en représentent 55.4 % (cf. Figure 3.15). Toute augmentation de la taille des échantillons de SNe ou de la qualité des données de suivi doit donc désormais s'accompagner d'un effort en vue de réduire la taille des incertitudes systématiques.

La taille de la statistique disponible dans le diagramme de Hubble va augmenter d'un facteur 7 d'ici 2025, avec la publication de l'échantillon ZTF (6000 SNe) à bas *redshift*, et des échantillons DES ($O(300^7)$ SNe à $z < 0.8$) et Subaru ($O(500)$ SNe à $z < 1.1 + 50$ SNe dans la gamme $1.1 < z < 1.5$). À plus long terme, les relevés LSST et Roman collecteront plusieurs dizaines de milliers de SNe supplémentaires.

La principale contribution au bilan des systématiques est la calibration photométrique du relevé. Celle-ci affecte non seulement les mesures photométriques, en particulier les rapports de flux

7. 300 SNe ont été confirmées Ia par spectroscopie, mais plus de 1500 passent les critères photométriques de candidats SN Ia [Vincenzi et al. (2022)].

entre bandes, mais aussi le modèle de courbe de lumière, car celui-ci est entraîné sur des données elles-mêmes entachées d'incertitudes de calibration. La calibration photométrique impacte les quantités reconstruites m_B^* et C et donc les distances via deux chemins : les données et le modèle. La communauté a engagé des efforts importants afin d'améliorer significativement les incertitudes de calibration photométriques. À ce jour, les chaînes de métrologie internes aux relevés garantissent une précision de l'ordre de quelque pour mille, entre standards de flux (CALSPEC) et supernovæ. Des expériences dédiées sont en cours pour contrôler la calibration des standards CALSPEC en les connectant à des standards de flux de laboratoire. Les sources suivantes d'incertitudes systématiques sont liées à notre connaissance imparfaite des SNe Ia. La première est la modélisation empirique des SNe. Du fait de la taille limitée de l'échantillon d'entraînement, le modèle de courbes de lumière est affecté d'un bruit qui se propage sur toutes les SNe. La seconde est issue de la convolution entre le biais de sélection des SN et les incertitudes astrophysiques sur le lien entre luminosité absolue des SNe et leur environnement. Pour attaquer les incertitudes de modélisation et les incertitudes astrophysiques, la clef est l'augmentation de la taille des échantillons, qui proviendra des relevés ZTF, LSST et Roman.

Mon travail de thèse est consacré au développement d'un nouveau cadre pour la modélisation des SNe Ia. Il s'agit de l'outil central pour l'inférence des distances aux supernovæ de type Ia à partir de leurs courbes de lumière. Le but de la thèse est de développer un cadre logiciel suffisamment rapide et léger pour pouvoir absorber les grandes statistiques à venir, et suffisamment versatile pour permettre de tester facilement de nouvelles paramétrisations de SNe Ia. Dans le prochain chapitre, nous retraçons brièvement l'histoire des techniques de modélisation empirique des SNe Ia. Les chapitres suivants seront consacrés à mon travail de thèse.

Deuxième partie

Développement d'une nouvelle génération de modélisation spectrophotométrique empirique des Supernovæ de type Ia : *Nouveaux algorithmes de Courbes de lumière (NaCl)*

Modèles empiriques de flux de Supernovæ de type Ia

Dans ce chapitre nous décrivons l’historique de la modélisation empirique des SNe Ia. Dans une première partie, nous rappelons l’importance de cette étape pour l’analyse de l’équation d’état de l’énergie noire. Dans la section suivante, l’approche historique de reconstruction de courbes de lumière dans le référentiel de la SN est introduit. Ensuite les trois principales modélisations de distribution en énergie spectrale sont présentées, dont le modèle SALT2. La section 4.4 présente trois modèles développés dans le cadre de l’analyse du relevé SNfactory. La section suivante présente deux ré-entraînements récents du modèle SALT2.

Contents

4.1	Introduction	102
4.2	Modèle d’ajustement de courbes de lumière	103
4.2.1	L’approche historique	103
4.2.2	MLCS/MLCS2k2	105
4.3	Modèle de distribution en énergie spectrale	105
4.3.1	SALT	106
4.3.2	SiFTO	106
4.3.3	SALT2	107
4.4	Les modèles de la SNfactory	109
4.4.1	SNEMO	109
4.4.2	SUGAR	111
4.4.3	Twins Embedding	111
4.5	Deux nouveaux ré-entraînements de SALT2	113
4.6	Conclusion	115

4.1 Introduction

Le signal cosmologique que nous cherchons à mesurer est encodé dans l'évolution relative du flux des supernovæ en fonction du *redshift*. Les flux à notre disposition sont des flux large bande, mesurés dans le référentiel observateur sur Terre. En raison du *redshift*, ceux-ci intègrent des zones du spectre qui diffèrent d'une SN à l'autre. Or, il est essentiel de comparer la même quantité physique pour tous les objets. On définit donc le flux comme le flux apparent correspondant aux photons émis dans une bande de référence, définie dans le référentiel de la SN.

En pratique, la bande B a été utilisée pour plusieurs raisons : c'est la zone spectrale où la luminosité des SNe Ia est maximale. C'est également la zone où la luminosité des SNe Ia est standardisable (la dispersion dans l'UV par exemple est beaucoup plus importante d'une SN à l'autre). Enfin, la longueur d'onde effective du filtre B étant $\lambda_B \sim 4350\text{\AA}$, cette définition permet de couvrir l'intégralité de la gamme en *redshift* $0 < z < 1.1$ avec des observations dans le visible et le proche infra-rouge, réalisées avec un détecteur silicium. Par la suite, on parlera de la bande B -restframe et on la notera B^* .

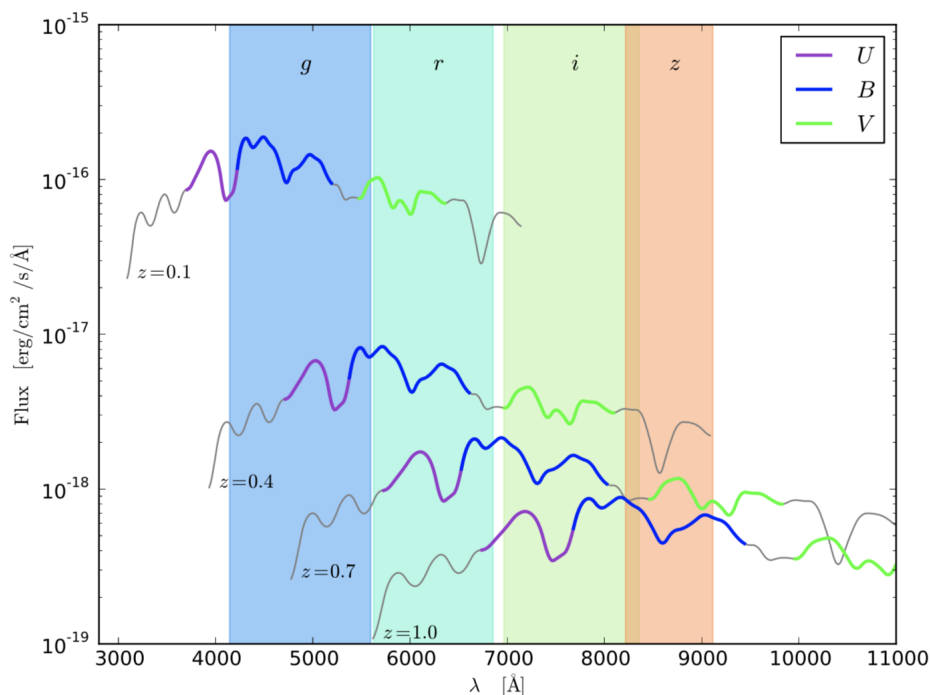


FIGURE 4.1 – Spectres de quatre SNe émis à la même phase et à des *redshift* de 0.1, 0.4, 0.7, et 1.0. Quatre bandes *griz* sont définies dans le référentiel de l'observateur. L'intégrale d'un spectre dans cette bande correspond à la magnitude observée. La partie surlignée en bleu de chaque spectre correspond au filtre B^* défini dans le référentiel de chaque SN. L'intégrale de cette partie de chaque spectre correspond à la quantité nécessaire pour inférer des distances. Pour un spectre émis à $z = 0.1$, cette quantité est observée en bande g . Pour un spectre émis à $z = 1$, elle est observée dans les bandes i , z .

Du fait du *redshift*, les flux B^* ne sont pas directement observables. La figure 4.1 illustre la situation : on souhaite reconstruire le flux correspondant à la zone bleue du spectre à partir des observations réalisées dans plusieurs bandes, ici les bandes g , r , i et z . Pour ce faire, on a besoin d'un modèle de l'évolution du spectre des supernovæ en fonction de leur phase.

On rappelle que la phase p , temps en jours comptés, dans le référentiel de la SN, à partir du maximum de luminosité dans la bande B^* défini dans le référentiel de la SN, s'écrit comme suit :

$$p = \frac{t - t_{max}^{B^*}}{1 + z} \quad (4.1)$$

t est le temps dans le référentiel de l'observateur, $t_{max}^{B^*}$ est le temps de maximum de luminosité dans la bande B^* , elle aussi dans le référentiel de l'observateur et z est le *redshift* de la SN.

À ce jour, les modèles théoriques d'explosion ne permettent pas de reproduire les observations avec une bonne précision. On utilise donc des modèles empiriques, construits à partir d'observations spectroscopiques et photométriques de SNe Ia. L'importance et la complexité de l'élaboration de ce type d'outil s'illustre simplement par le nombre de différents développements qui en ont été réalisés et dont il est dressé ici, une liste non exhaustive : MLCS, stretch, Δm_{15} , superstretch, BATM, SALT, CMAGIC, MLCS2k2, SALT2, SiFTO, SNooPy, SNEMO, SUGAR, Twin embedding . . .

En pratique on constate que l'évolution spectrale des SNe Ia varie d'un objet à l'autre. Pour être précis, les modèles doivent donc tenir compte de la diversité spectrale et photométrique des supernovæ. Empiriquement, la population semble être correctement décrite avec deux paramètres. Cette modélisation est guidée par le lot de données sur lesquelles est entraîné le modèle. Sa finesse est fonction du nombre d'objets du lot d'entraînement et de la qualité des observations.

La plupart des modèles modernes sont entraînés simultanément sur des observations spectroscopiques et photométriques. Le modèle une fois entraîné est une fonction de la phase, du *redshift*, de la date de maximum, de la magnitude exprimée dans le référentiel de la SN et de paramètres décrivant la diversité des SNe (typiquement un paramètre de forme comme le *stretch*, et une couleur elle aussi calculée à partir de courbes de lumière exprimée dans le référentiel de la SN). Il est ensuite ajusté, pour une SN donnée, sur ses courbes de lumière mesurées ce qui permet de déterminer ses paramètres d'intérêt que sont la magnitude B^* , le paramètre de forme et la couleur.

Dans la suite de cette section, nous donnons un bref aperçu de l'évolution des modèles de courbes de lumière au cours des deux dernières décennies.

4.2 Modèle d'ajustement de courbes de lumière

4.2.1 L'approche historique

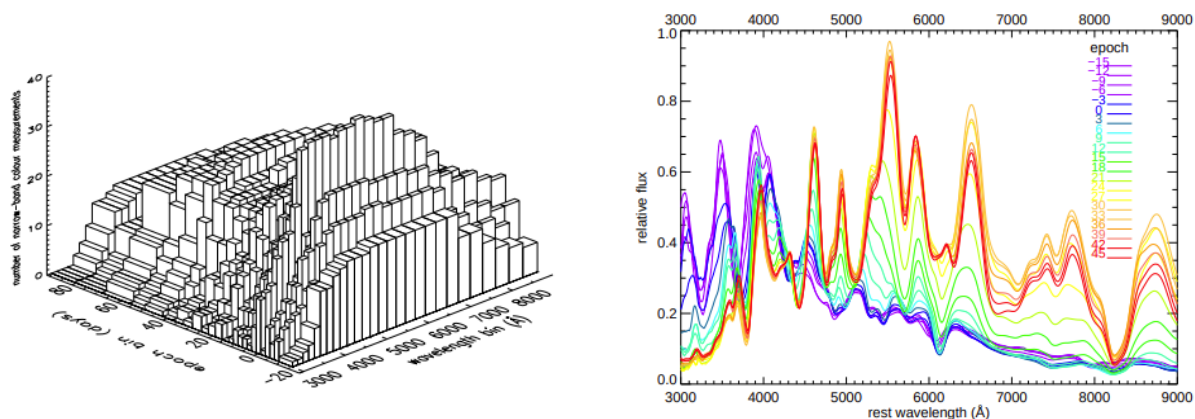
L'approche historique ne fait pas intervenir un modèle proprement dit, mais plutôt une série de transformations des données. Il s'agit tout d'abord d'interpoler en longueur d'onde pour déterminer à chaque phase, à partir de mesures de flux dans différentes bandes, une courbe de lumière B^* (correction- K). Puis d'interpoler, à l'aide de splines ou d'une fonction analytique entre les points de la courbe de lumière ainsi reconstruite, afin de déterminer la magnitude B^* au pic.

La K -correction Pour corriger l'effet du *redshift* sur le spectre (cf. Figure 4.1), on ajoute un *offset* à la magnitude observée noté K et appelé *K -correction* [Humason et al. (1956); Oke &

Sandage (1968); Hamuy et al. (1993b)]. La magnitude observée dans la bande X, exprimée dans le référentiel de la SN est définie par :

$$m_X(z=0) = m_X(z) - K(z) \quad \text{avec} \quad K(z) = 2.5 \log \left((1+z) \frac{\int S(\lambda) T_X(\lambda) d\lambda}{\int S(\lambda/(1+z)) T_X(\lambda) d\lambda} \right) \quad (4.2)$$

$m_X(z)$ est la magnitude observée, z le *redshift* de la SN et S est le modèle de la distribution spectrale en énergie évalué à la même phase. Historiquement, la correction K était calculée en utilisant le spectre d'une supernova observée à une phase similaire à l'observation $m_X(z)$. Ensuite des modèles du spectre en fonction de la phase de la SN ont été assemblés [Nugent et al. (2002); Hsiao et al. (2007)] à partir de spectres de supernovæ proches, recalibrés, car mesurés par différents instruments, et interpolés créant une surface spectrale $S(\lambda, p)$ continue et à haut signal sur bruit.



(a) Surface spectrale réalisée avec ~ 600 spectres (recalibrés et interpolés) de ~ 100 SNe, à des *redshifts* jusqu'à 0.8, et des phases comprises entre -20 et 100 jours. Figure extraite de Hsiao et al. (2007). (b) Modélisation de 22 spectres dont les phases sont entre -15 et 45 jours, coupes de la surface spectrale (Figure de gauche).

FIGURE 4.2 – Surface spectrale et coupes à phase fixée. Figure extraite de Hsiao et al. (2007)

Modèles de courbes de lumière Une fois les données " K -corrigées", on ajuste un modèle pour déterminer la magnitude B^* au pic et extraire les paramètres de standardisation. Les distances sont directement inférées de ce processus. Historiquement, ces modèles ne paramétrisaient la variabilité qu'avec un paramètre de forme. La variation en couleur était supposée due exclusivement à l'absorption de la lumière par les poussières présentes sur la ligne de visée, et non à une variabilité intrinsèque. La paramétrisation de cette extinction suit une loi de Cardelli [Cardelli et al. (1989)], modélisant l'absorption due aux poussières à la longueur d'onde λ , A_λ , par la relation suivante :

$$m_\lambda - m_{\lambda_0} = A_\lambda = A_V \left(a_\lambda + \frac{b_\lambda}{R_V} \right) \quad (4.3)$$

m_λ est la magnitude observée et m_{λ_0} est la magnitude qui serait observée en l'absence d'extinction. a_λ et b_λ sont des polynômes déterminés empiriquement. Les deux paramètres A_V et R_V correspondent respectivement à l'absorption sur la ligne de visée définie dans la bande V et à un

ratio dépendant des propriétés de la poussière. Ils sont reliés à l'aide de l'excès de couleur entre les bandes B et V , $E(B - V)$:

$$A_V = E(B - V) \times R_V \quad (4.4)$$

Le ratio R_V a une valeur moyenne de 3.1 pour l'absorption due à la poussière de la Voie Lactée.

4.2.2 MLCS/MLCS2k2

Le modèle MLCS2k2 (acronyme pour *Multicolor Light-Curve Shapes 2002*) a été développé pour l'analyse du suivi de la *High-Z Supernova Search Team* (Riess et al. (1996), Jha et al. (2007)). La magnitude dans la bande X observée à une phase p est modélisée par la formule suivante :

$$m_X(p) = \mu + Ext_X^{VL} + Ext_{X'}^{gal} + K_{XX'}^{cor} + M_{X'}(p) + P_{X'}(p)\Delta + Q_{X'}(p)\Delta^2 \quad (4.5)$$

où μ est le module de distance, Ext_X^{VL} est l'extinction de la Voie Lactée suivant une loi de Cardelli, corrigée dans le filtre du référentiel de l'observateur, $Ext_{X'}^{gal}$ est l'extinction due à la galaxie hôte comptée dans le référentiel de la SN, suivant là aussi une loi de Cardelli à deux paramètres : A_V et R_V ¹. $K_{XX'}^{cor}$ est le terme de K -correction entre le filtre d'observation X et le filtre de référence exprimé dans le référentiel de la SN $X' = U^*, B^*, V^*, R^*$ ou I^* . Δ est le paramètre de forme temporelle, tenant compte de la variabilité des SNe. Enfin $M_{X'}$, $P_{X'}$, $Q_{X'}$ sont les fonctions qui décrivent les courbes de lumière de référence et qui ont été extraites d'un entraînement réalisé sur environ 100 SNe à bas *redshift*. $M_{X'}$ correspond donc à la magnitude absolue lorsque $\Delta = 0$ et $Q_{X'}$ et $P_{X'}$ sont des corrections du premier et du second ordre en Δ .

Pour chaque SN, le terme de K -correction est déterminé avant l'ajustement et ses paramètres ensuite ajustés sont donc Δ , $t_{max}^{B^*}$, μ et A_V . La variation de couleur est exclusivement attribuée à l'extinction causée par la galaxie hôte et par celle de la Voie Lactée.

Un présupposé important de cette paramétrisation est donc que la description de la variabilité des SNe ne nécessite pas de paramètres de couleur. Cette paramétrisation implique aussi que la totalité de la variabilité des SNe n'est pas prise en compte par le modèle, car sans description intrinsèque. MLCS2k2 fournit également un modèle d'erreur sur le modèle d'ajustement des courbes de lumière, permettant de prendre en compte la variabilité intrinsèque non décrite par le modèle a posteriori, en ajoutant une paramétrisation de la variance qui lors d'un ajustement par moindre carré permet d'obtenir un χ^2 par degrés de liberté égale à 1.

4.3 Modèle de distribution en énergie spectrale

Ajuster un modèle sur une fonction des données plutôt que sur les données elles-mêmes complique considérablement la propagation des erreurs. Une évolution naturelle consiste donc à développer un modèle prédisant directement une courbe de lumière dans une bande donnée, en intégrant l'étape de K -correction. Ce type de modèle est construit à partir d'une surface spectrale $S(\lambda, p)$ que l'on s'autorise à déformer puis à intégrer dans un modèle de la bande passante dans laquelle l'observation est effectuée.

1. Pouvant être fixé a 3.1 [Jha et al. (2007)] ou estimé lors de l'ajustement [Kessler et al. (2009)].

4.3.1 SALT

Le modèle SALT (*Spectral Adaptive Light curve Template*, Guy et al. (2005)) a été développé dans le cadre de l'analyse du relevé SNLS. Il est construit à partir de la surface spectrale déterminée par Nugent et al. (2002), déformée par deux fonctions polynomiales, l'une décrivant l'effet du *stretch* s et l'autre de la couleur c . Un point de photométrie, d'une SN de *redshift* z , observé à la phase $p_s = \frac{t-t_{max}^*}{s(1+z)}$, à travers un instrument de transmission T_X est modélisé par la formule suivante :

$$F_X(t) = f_0(1+z) \int S(p_s, \lambda) \frac{\lambda}{hc} T_X(\lambda(1+z)) d\lambda e^{-0.4 \ln(10) \mathcal{K}(p_s, \lambda_{T_X}, s, c)} \quad (4.6)$$

où f_0 est le facteur de normalisation directement relié à m_{B^*} , S est la surface spectrale modélisée par Nugent et al. (2002), s et c sont le *stretch* et la couleur de la SN. \mathcal{K} est séparé en deux fonctions polynomiales, une dépendante du *stretch* et une dépendante de la couleur :

$$\mathcal{K}(p_s, \lambda_{T_X}, s, c) = \mathcal{K}_s(p_s, \lambda_{T_X}, s) + \mathcal{K}_c(\lambda_{T_X}, c) \quad (4.7)$$

\mathcal{K}_s est un polynôme de degré 4 en phase, 3 en longueur d'onde et 1 en *stretch*. \mathcal{K}_c est un polynôme de degré 3 en longueur d'onde et 1 en couleur. Il y a donc au total 48 paramètres à déterminer mais en pratique seulement 34 car des contraintes sont imposées au modèle pour briser les dégénérescences. L'ajustement de ces paramètres est réalisé sur des supernovæ à bas *redshift* et constitue l'entraînement du modèle. Le modèle entraîné est ensuite ajusté sur les courbes de lumière des SN étudiées, les paramètres d'intérêt de l'ajustement étant la normalisation f_0 , le *stretch* s et la couleur c de la SN. Ces paramètres sont ensuite combinés lors du fit du diagramme de Hubble, comme décrit au chapitre précédent.

Ce modèle a été le premier à implémenter une paramétrisation globale à toutes les SNe, expliquant leur variabilité en couleur par une contribution intrinsèque à la SN. Sa spécificité est de paramétrer la variabilité en couleur de manière globale, avec une loi commune à toutes les SNe, sans faire aucune hypothèse sur l'origine physique de cette variabilité.

4.3.2 SiFTO

Le modèle présenté dans Conley et al. (2008), appelé SiFTO, est également issu de l'effort d'analyse des données de SNLS². Comme SALT, SiFTO prédit des flux photométriques dans des bandes définies dans le référentiel de l'observateur et peut être ajusté sur des courbes de lumière. Comme SALT, il utilise également comme à priori un modèle de distribution d'énergie spectrale détaillé dans Hsiao et al. (2007) (Figure 4.2). Cette surface est construite de manière à conserver les amplitudes relatives de caractéristiques spectrales consécutives. Les variations sur de larges intervalles de longueur d'onde, entre les différentes bandes d'observation ne sont pas prises en compte. Une étape préliminaire pour corriger ces effets sur des données de SNe à petit z et de SNLS, est donc réalisée.

Le modèle est une paramétrisation à deux paramètres du flux de la supernova qui sont directement liés aux paramètres de standardisation des Supernovæ, le *stretch* et la couleur. SiFTO définit le

2. Pour l'analyse de ce relevé tous les outils ont été développés en parallèle par les équipes française et canadienne, permettant d'obtenir deux analyses indépendantes.

stretch dans la bande B , s_B . Le *stretch* peut varier en fonction de la longueur d'onde *restframe* comme une fonction de s_B . La couleur est quant à elle, paramétrisée de manière différente. C'est la normalisation de l'ajustement des courbes de lumière, indépendante entre les différentes bandes. Le flux photométrique émis à un *redshift* z et observé à un temps t à travers un instrument X de transmission T_X est défini par :

$$F_X(t) = \frac{1}{(1+z) \int T_X(\lambda) \lambda d\lambda} \int \phi_0 \left(\frac{\lambda}{1+z}, p_{s_B} \right) S \left(\frac{\lambda}{1+z}, n_X R_X(p_{s_B}, p_{s_X}, z) \right) T_X(\lambda) \lambda d\lambda \quad (4.8)$$

Le modèle de référence ϕ_0 est corrigé d'un facteur de normalisation pour obtenir la couleur observée dans le référentiel de l'observateur S . C'est une fonction de splines, uniquement déterminée par le produit $n_X R_X$, produit de l'amplitude du filtre correspondant, et du rapport de normalisation, fonction du *stretch* de la SN, et défini par :

$$R_X(t, t_0, s_B, s_X, z) = \frac{\int \phi_0 \left(\frac{\lambda}{1+z}, \frac{t-t_0}{s_X(1+z)} \right) T_X(\lambda) \lambda d\lambda}{\int \phi_0 \left(\frac{\lambda}{1+z}, \frac{t-t_0}{s_B(1+z)} \right) T_X(\lambda) \lambda d\lambda} \quad (4.9)$$

S est ajusté pour que la SED ait la couleur observée. L'entraînement du modèle a été réalisé sur 72 SNe à bas *redshift* provenant de différents relevés et 98 SNe observées à haut *redshift* par l'expérience SNLS. On peut voir sur la figure 4.3, les différents composants du modèle SiFTO, appliqué à la SN 06D2ag. La paramétrisation de SiFTO ne permet pas de décrire toute la variabilité des SNe et un modèle d'erreur est ajusté aux données.

4.3.3 SALT2

Le modèle SALT2 [Guy et al. (2007)] pousse plus avant la modélisation en intégrant la construction de la surface spectrale dans l'entraînement du modèle, et en tentant de décrire la variabilité des SNe au niveau de la surface spectrale elle-même. L'idée initiale était de réaliser une analyse en composantes principales (PCA) des données spectrales et photométriques disponibles, dans l'espace phase-longueur d'onde (*restframe*). Dans cette approche, les composantes principales sont des surfaces, décomposées sur des bases de splines 2D. En pratique, la taille limitée de l'échantillon d'entraînement a conduit à tronquer la PCA à l'ordre 1, et se contenter d'une loi de couleur globale (indépendante de la phase) :

$$S(\lambda, t) = X_0[M_0(\lambda, p) + X_1 M_1(\lambda, p)] e^{CL(\lambda)C} \quad (4.10)$$

M_0 et M_1 sont les deux surfaces qui décrivent la distribution spectrale moyenne et la première variation. $CL(\lambda)$ est la loi de couleur des SNe. Elle est complètement empirique (implémentée comme un polynôme) et ne fait aucun a priori sur la source de variabilité en couleur. Elle peut donc prendre en compte des phénomènes extrinsèques (provenant de l'absorption des poussières sur la ligne de visée) et intrinsèques (propres à la SN). À ceux-ci s'ajoutent les paramètres propres

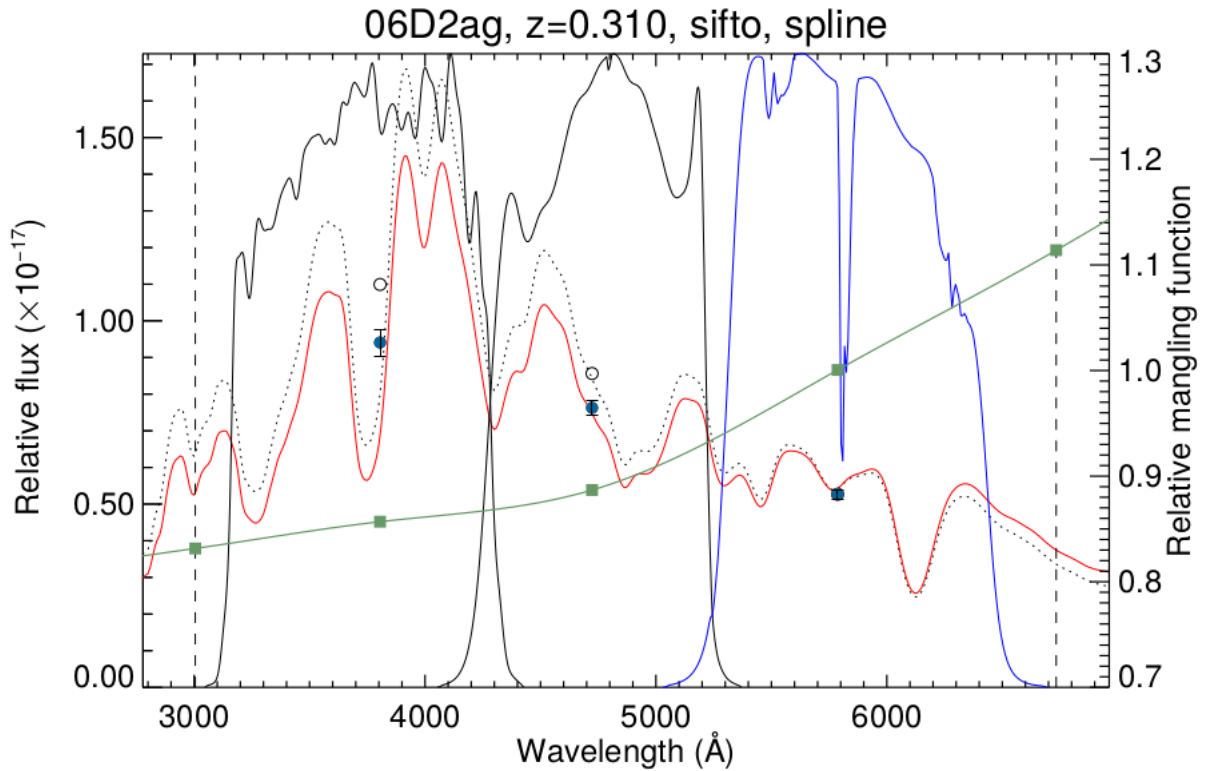


FIGURE 4.3 – Ingrédients du modèles SiFTO évalué pour trois points de photométrie de la SN 06D2ag. En arrière plan, les transmissions des trois filtres (à ce *redshift* $z \sim 0.31$, l'information contenue dans bande z est hors des limites du domaine de définition du modèle) de MegaCam, g , r et i . Les points de photométrie sont représentés avec les barres d'erreur en noir, en tirets, le spectre, extrait du modèle de référence ϕ_0 . Les cercles creux noirs sont le flux intégré de ce spectre dans les différentes bandes. En rouge, la SED ajustée en couleur, grâce à la fonction de splines S (en vert) et dont l'intégrale à travers chaque bande donne les points de photométrie (cercles bleus).

à chaque SN : X_0 est la normalisation du modèle, X_1 est un paramètre de forme analogue au stretch et C une couleur restframe. En pratique, le modèle peut-être défini de façon à ce que X_0 puisse être identifié à l'amplitude de la courbe de lumière B^* , X_1 au stretch et C à un paramètre de couleur. En pratique, le modèle SALT2 ne réalise pas de K-correction en bande B^* mais utilise une amplitude globale (Offset) pour définir : $m_{B^*} = \text{Offset} - 2.5 \log(X_0)$.

À partir de ce modèle de SED, il est possible d'exprimer le flux photométrique d'une SN émis à une phase p , observé à travers un filtre X , de transmission connu T_X par :

$$F_X(t) = \int S(\lambda, t) T_X(\lambda(1+z)) \frac{\lambda}{hc} d\lambda \quad (4.11)$$

L'entraînement du modèle consiste à déterminer, par un ajustement par moindres carrés sur les spectres et les courbes de lumière des SNe du lot d'entraînement, les deux surfaces M_0, M_1 , la loi de couleur, ainsi que les paramètres $(X_0, X_1, t_{max}^{B^*}, C)$ propres à chaque SN.

L'entraînement de ce modèle a été effectué plusieurs fois au fur et à mesure que la statistique disponible augmentait [Guy et al. (2007, 2010); Betoule et al. (2014)]. Il peut être réalisé sur des SNe à bas et haut *redshift* puisqu'il ne calcule pas de distances directement. Cela permet au modèle d'être représentatif de la population moyenne du diagramme de Hubble. Cela permet également de contraindre le modèle dans l'UV proche, grâce aux bandes *g* et *r* des SNe à haut *redshift*, sans pour autant avoir recours aux données *U* sol, notoirement difficiles à calibrer.

L'un des derniers entraînement de SALT2, noté SALT2.4, a été réalisé dans le cadre de la *Joint Light curve Analysis* (JLA). Du fait de sa simplicité de modélisation, sa gestion de la loi de couleur (sans à priori) et sa grande flexibilité pouvant être entraîné à la fois sur les courbes de lumière et de spectres, le modèle SALT2 représente aujourd'hui l'état de l'art de la modélisation empirique du flux de SNe et est utilisé³ dans les analyses cosmologiques récentes [Betoule et al. (2014); Scolnic et al. (2018); Riess et al. (2021); Brout et al. (2022)].

4.4 Les modèles de la SNfactory

Une troisième génération de modèles, comprenant les modèles SNEMO [(Saunders et al. 2018)], SUGAR [(Léget et al. 2020)] et *Twins Embedding* [(Boone 2021)], a été développée sur les séries spectrophotométriques des SNe Ia proches découvertes et suivies par la SNfactory entre 2004 et 2014. Le lot de données produit par SNfactory est le seul qui donne accès aux variations fines des propriétés spectrales des SNe Ia sur une large plage de phase. C'est donc un lot d'entraînement unique pour un modèle spectrophotométrique tel que SALT2. En revanche, ce type de séquences spectrales n'est aujourd'hui disponible que dans la gamme de *redshift* $z < 0.15$. Il est donc impossible de généraliser l'approche dans la gamme $0.05 < z < 1.1$ utilisée pour la mesure de w .

Les modèles développés au sein de la SNfactory se concentrent donc sur la modélisation de la diversité des SNe Ia.

4.4.1 SNEMO

SNEMO (pour *SuperNova Empirical MOdels*), décrit dans Saunders et al. (2018), est une généralisation de la technique de PCA développée pour SALT2. Les composantes principales sont des surfaces spectrales $S(\lambda, p)$. Elles sont extraites directement des séries spectrales en utilisant une EMFA (pour *Expectation Minimization for Factor Analysis*). L'EMFA est une méthode équivalente à la PCA, qui prend en compte les erreurs de mesure affectant les données analysées.

Plusieurs entraînements sont générés en poussant l'EMFA à différents ordres. Ces composantes principales forment une base linéaire décrivant n'importe quelle série temporelle spectrale en optique f_{SN} :

$$f_{SN} = c_{SN,0} \left(e_0(\lambda, p) + \sum_{i=1}^N c'_{SN,i} e_i(\lambda, p) \right) \quad (4.12)$$

Les $c'_{SN,i} = \frac{c_{SN,i}}{c_{SN,0}}$ sont normalisés afin de n'avoir qu'un seul paramètre dépendant de la distance de luminosité, $c_{SN,0}$. Dans cette équation, F_{SN} est le flux de la supernova émis à la phase p et la longueur d'onde λ . N est le nombre de composantes du modèle, c_{SN} sont les coefficients

3. Ou ré-entraînement.

du modèle (propre à chaque SN) et e_i sont les composantes du modèle (les vecteurs propres, commun à toutes les SNe). Le rôle de l'entraînement est de déterminer ces vecteurs propres.

Afin de séparer la variabilité en couleur intrinsèque aux SNe et la variabilité en couleur due à l'absorption par les poussières, SNEMO tente de corriger explicitement les spectres de l'absorption sur la ligne de visée avant entraînement. Pour cela, une loi d'absorption du type $10^{-0.4A_s C(\lambda)}$ est ajustée sur chaque spectre, où $C(\lambda)$ est la loi d'absorption proposée par Cardelli et al. (1989) pour décrire l'absorption par les poussières galactiques (R_V est fixé à 3.1) et A_s est l'extinction de chaque SN, par rapport à une surface fiducielle. En pratique, l'ajustement est réalisé simultanément sur tout le lot de données, afin de pouvoir ajuster la surface fiducielle conjointement avec les lois d'absorption de chaque SN. Chaque spectre est ensuite dérougi par la loi de rougissement qui lui correspond :

$$F_{SN}(\lambda, p) = f_{SN}(\lambda, p) 10^{0.4A_s C(\lambda)} \quad (4.13)$$

Dans Saunders et al. (2018) trois modèles sont présentés : SNEMO2 avec deux surfaces permet une comparaison avec SALT2 (cf. Figure 4.4). On constate effectivement une corrélation entre les paramètres de SNEMO2 et les paramètres x_1 et c de SALT2. SNEMO7 avec 7 surfaces décrit davantage la variabilité des SNe. Enfin SNEMO15 décrit au mieux la variabilité spectrale de l'échantillon sans être sur-entraîné.

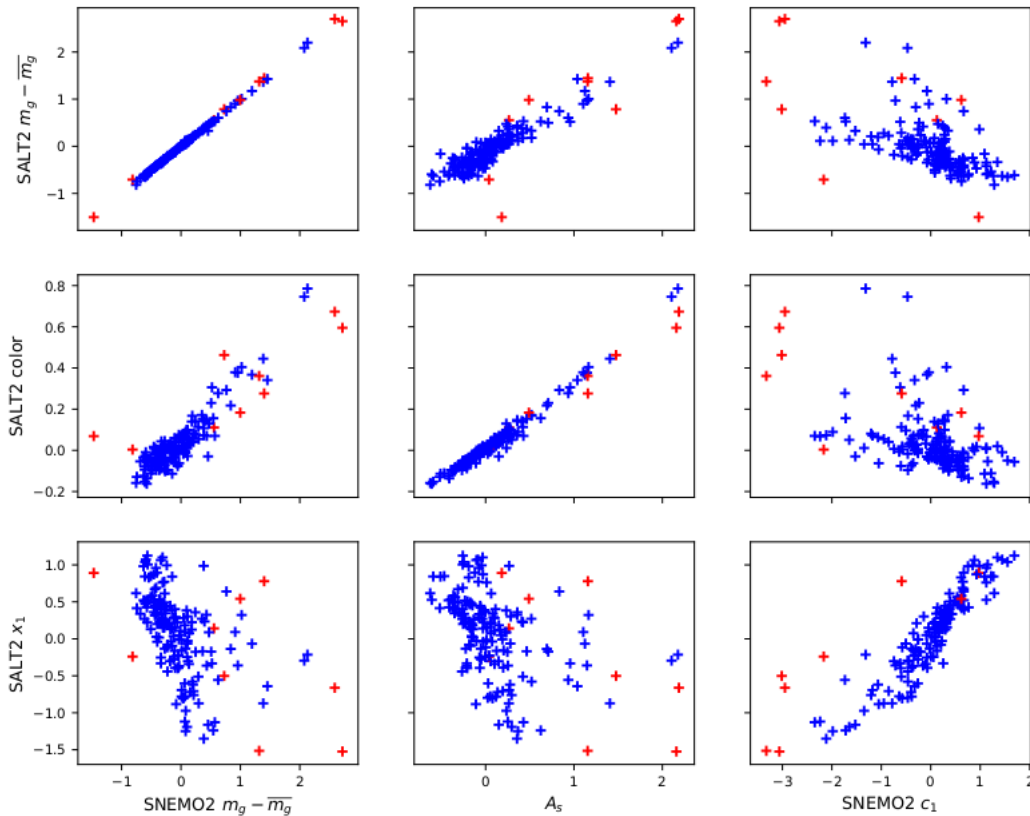


FIGURE 4.4 – Corrélation entre les paramètres de SNEMO2 et SALT2, extraits sur le même lot de données. Les points rouges correspondent aux SNe définies comme particulières. Figure extraite de Saunders et al. (2018).

4.4.2 SUGAR

SUGAR, (pour *SUPernova Generator And Reconstructor*) est décrit dans Léget et al. (2020). Comme SNEMO, le modèle est basé sur un algorithme d'analyse en facteurs (EMFA), mais réalisé dans un espace différent : l'idée est que l'essentiel de la diversité des SNe est visible dans leur caractéristiques spectrales (largeur équivalente ou minima du profil P-Cygni d'une sélection de raies) des spectres proches du maximum de luminosité. Les concepteurs de SUGAR ont sélectionné 13 caractéristiques spectrales encodant l'essentiel de la variabilité de SN à SN.

La première partie de l'algorithme consiste à extraire ces caractéristiques spectrales des spectres proches du maximum, et à réaliser une analyse EMFA dans cet espace. L'analyse montre que l'essentiel de la variance peut-être capturé par 3 composantes. Les deux premières sont corrélées aux profils *P-Cygni* des principaux éléments présents dans les spectres, et aux vitesses des raies d'absorption. Le modèle proprement dit est ensuite construit, en ajustant les surfaces moyennes correspondant à chaque composante. Au préalable, les spectres sont dérougés, en suivant une procédure analogue à celle mis en place dans le modèle SNEMO.

Le modèle SUGAR est défini comme :

$$M_i(\lambda, t) = M_0(\lambda, t) + q_i^j \alpha_j(\lambda, t) + A_{V,i} f(\lambda, R_V) + \Delta M_i^{grey} \quad (4.14)$$

où M_0 est la surface spectrale moyenne, les surfaces α_j associées aux valeurs propres q_i^j de l'EMFA. $A_{V,i}$ est l'extinction en bande V , f est la loi d'extinction de Cardelli, le ratio R_V étant déterminé à la seconde étape de création du modèle et enfin ΔM_i^{grey} est une normalisation indépendante du temps. Les paramètres propres à chaque SN sont les paramètres q , A_V et ΔM^{grey} .

Les paramètres q_1 et q_3 de SUGAR sont corrélés aux paramètres X_1 de SALT2, le paramètre d'absorption, paramétrisant la loi d'extinction A_V est quand à lui corrélé à la couleur c de SALT2.

4.4.3 Twins Embedding

La technique est décrite dans Boone (2021). Ce modèle construit une paramétrisation non-linéaire de la diversité intrinsèque des SNe Ia, sur des supernovæ dites "jumelles". Dans Fakhouri et al. (2015) une nouvelle méthode de standardisation des SNe est introduite, reposant sur la distance spectrale d_{ij}^s entre deux SNe. Les SNe ayant une distance spectrale basse sont appelées "jumelles". Ces deux études ont elles aussi été réalisées avec les données de la *Nearby Supernovæ Factory*. Le modèle est construit en trois étapes. D'abord, une modélisation de la surface spectrale proche du maximum de luminosité est réalisée pour estimer le spectre moyen au maximum de luminosité des SNe, f_{mean} . Cette estimation du spectre moyen permet de prendre en compte les deux principaux phénomènes extrinsèques que sont l'incertitude sur le *redshift* de la galaxie hôte et les effets d'extinction dus à la poussières. La première contribution a pour effet d'ajouter un terme constant en magnitude. Ce terme est ajusté avec le modèle. Pour prendre en compte l'extinction, une loi de Cardelli décrite dans Chotard et al. (2011), de paramètres $R_V = 2.8 \pm 0.3$ est ajustée simultanément.

Ces spectres étant corrigés des phénomènes extrinsèques, la variabilité résiduelle peut être paramétrisée par un modèle non-linéaire. Ce modèle est ajusté par un algorithme de réduction de dimensionnalité non-linéaire, appelé carte iso ou (Isomap). C'est une généralisation de la PCA pour être sensible aux structures non-linéaires des données. Il est basé sur la recherche de plus

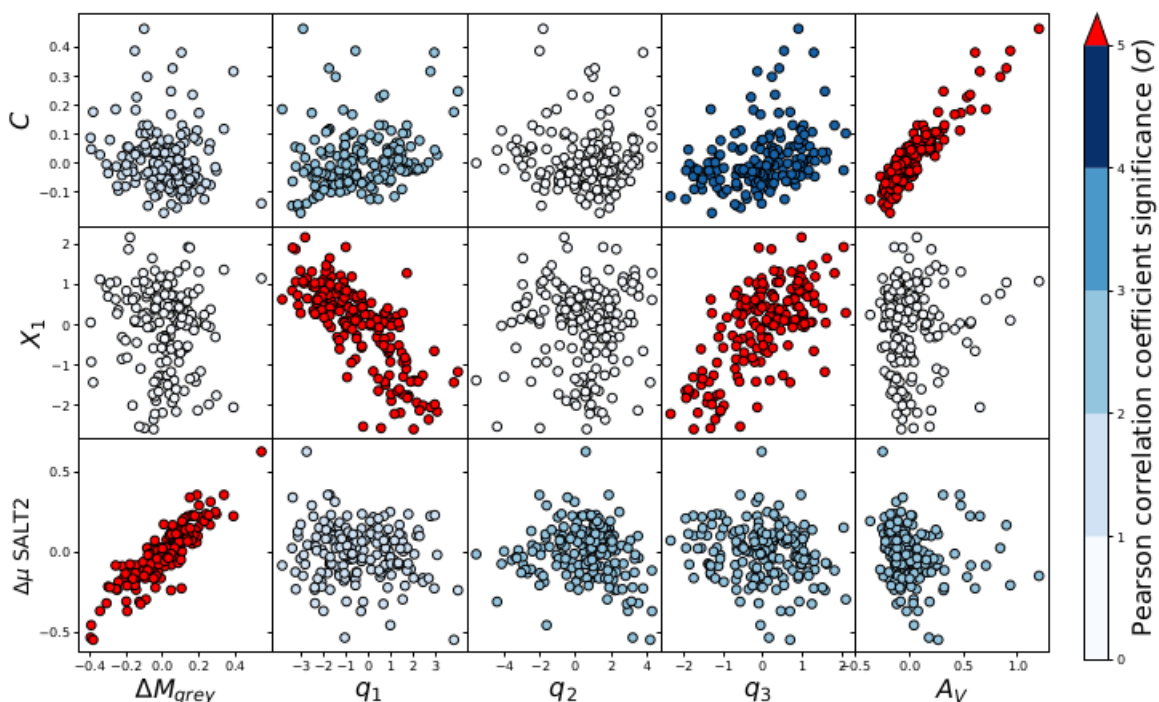


FIGURE 4.5 – Coefficients de corrélation entre les paramètres de SUGAR et SALT2. Figure extraite de Léget et al. (2020).

proche voisin défini selon une certaine distance. Elle est définie comme la distance spectrale définissant les SNe "jumelles", d_{ij}^s entre deux SNe, i et j , de spectres préalablement dérougés f_i et f_j :

$$d_{ij}^s = \left[\sum_k \left(\frac{f_i(\lambda_k) - f_j(\lambda_k)}{f_{mean}(\lambda_k)} \right)^2 \right]^{\frac{1}{2}} \quad (4.15)$$

Les paramètres du modèle sont la dimension de l'espace d'arrivée et le nombre de voisins. Ils sont définis comme les SNe dont la distance spectrale est la plus proche de la SN. Il est montré que le nombre de voisins n'a pas un grand impact sur la modélisation. Ce procédé ne permet pas d'avoir un modèle pour le spectre d'une SN suivant ces coordonnées du modèle. Des processus gaussiens sont utilisés pour modéliser un spectre en fonction des paramètres de la carte iso, précédemment déterminés (cf. Figure 4.6). Le modèle capturant le plus grande variabilité des SNe possède 3 composantes (cf. Figure 4.7).

La figure 4.8 présente les coefficients de corrélation entre les trois paramètres du modèle : ξ_1 , ξ_2 et ξ_3 et les principales caractéristiques spectrales des SN Ia, des paramètres de SALT2, SNEMO et SUGAR. Ces paramètres sont corrélés aux principaux indicateurs de variabilité des SNe Ia. Le paramètre décrivant la plus grande variabilité, ξ_1 , est fortement corrélé aux paramètres c_1 de SNEMO7, q_2 de SUGAR et aux pics du silicium et du calcium. Ce paramètre q_2 de SUGAR n'est lui-même pas corrélé aux paramètres de SALT2 (cf. Figure 4.5). Le paramètre ξ_2 est corrélé au paramètre X_1 , aux paramètres de SNEMO7 c_2 , c_3 et au paramètres de SUGAR q_1 et q_3 . ξ_3 est corrélé aux paramètres de SNEMO7 non corrélés aux paramètres précédents et à la vitesse du silicium.

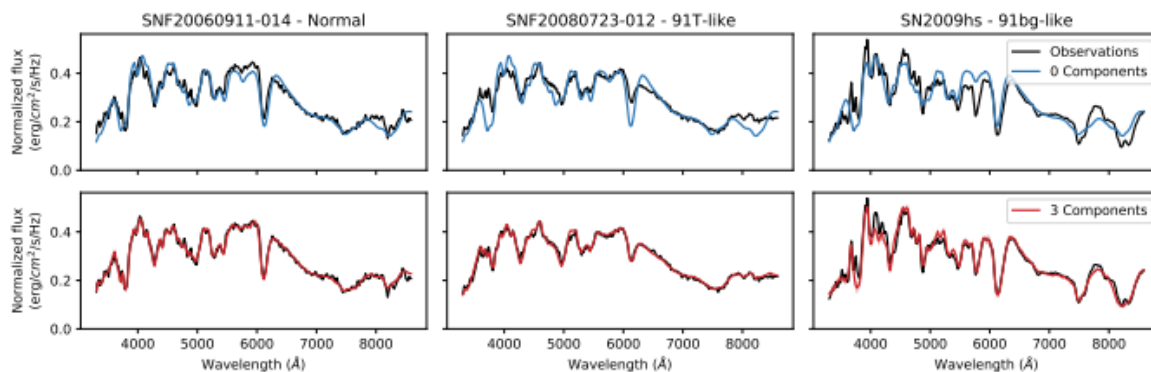


FIGURE 4.6 – Spectres de trois SNe Ia appartenant à trois sous-classes différentes : normale à gauche, 91T-like au milieu et 91-bg-like à droite. Les courbes bleues représentent le modèle issu des deux premières étapes sans paramétrisation non-linéaire de la variabilité intrinsèque. Les courbes rouges correspondent à un modèle de carte iso possédant trois paramètres et ajusté par processus gaussiens. Figure adapté de Boone (2021)

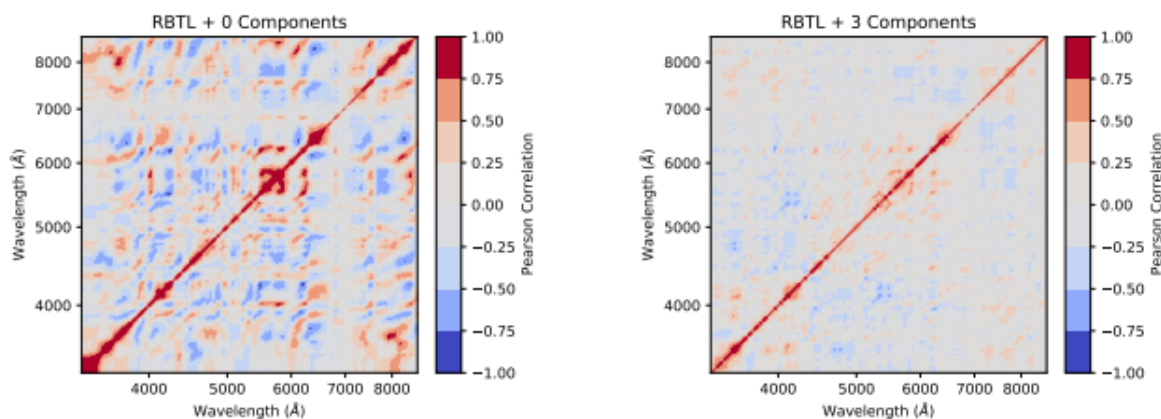


FIGURE 4.7 – Matrices de covariance des résidus des spectres de SNe Ia pour deux modèles. À gauche le modèle estimé à la première étape corrigé des effets extrinsèques. À droite le modèle de carte iso avec trois paramètres. Figure extraite de Boone (2021).

4.5 Deux nouveaux ré-entraînements de SALT2

Récemment deux nouveaux entraînements de SALT2 ont été publiés.

SALT2-2021 Il s’agit d’un re-entraînement à partir du lot JLA original, en utilisant le code historique de SALT2 (Guy et al. 2007). Il a été publié par Taylor et al. (2021) et utilisé dans l’analyse cosmologique Pantheon (Brout et al. 2022).

Les principales différences avec le modèle utilisé pour l’analyse JLA Betoule et al. (2014) sont l’utilisation de la carte d’extinction par les poussières Galactiques mise à jour (Schlafly & Finkbeiner 2011), ce qui conduit à une diminution de l’excès de couleur moyen (i.e. colonne de poussière) affectant la lumière des SNe d’environ 14% ; la modification de la calibration photométrique du lot d’entraînement, suite au travail mené par Scolnic et al. (2015) (*Supercal*), qui, en comparant les magnitudes des étoiles de champ, a réaligné les principaux échantillons de SNe disponibles

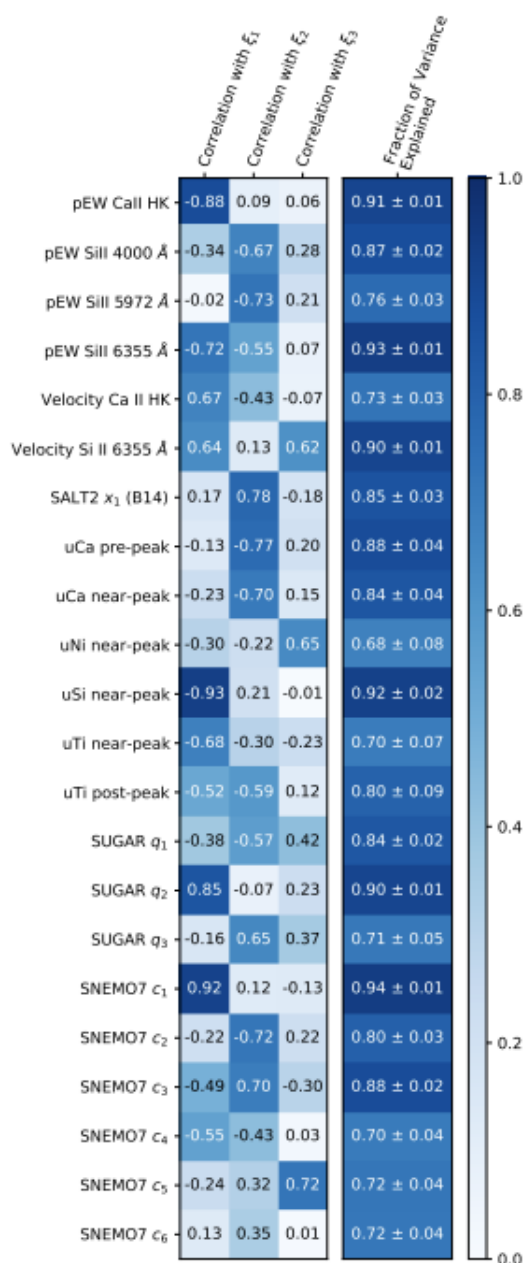


FIGURE 4.8 – Coefficients de corrélation entre les trois paramètres du modèles (colonnes) en fonction de caractéristiques spectrales et des paramètres de SALT2, SNEMO et SUGAR (lignes). Figure extraite de Boone (2021)

sur le catalogue PanSTARRS (PS1). Un changement de calibration du lot d’entraînement affecte le modèle moyen (surface M_0 et loi de couleur principalement) ainsi que le bilan d’incertitude du fait de l’impact de la calibration sur le modèle. Les nouvelles incertitudes de calibration ont été propagées dans l’entraînement en suivant la procédure explicitée dans Betoule et al. (2014).

SALT3 En 2021, un nouveau code d’entraînement du modèle SALT2, SALTShaker ainsi qu’un lot d’entraînement plus important que celui de JLA ont été mis à disposition par Kenworthy et al. (2021) (K21). Le lot d’entraînement, que nous appellerons K21, comprend l’échantillon original JLA, augmenté des données PanSTARRS-1 (PS1), DES et Panthéon (cf. Figure 4.9). Le code

d'entraînement entièrement écrit en python est public. La famille de modèles résultant est appelée SALT3.

La paramétrisation du modèle SALT3 est extrêmement proche de celle de SALT2. Les différences sont anecdotiques. Elles portent sur le schéma de régularisation du modèle, sur le fait que le modèle d'erreur n'est ajusté que sur la photométrie, et enfin, sur le fait que la loi de couleur n'est pas présente dans la partie spectrale du modèle – ce qui n'a aucun impact du fait de présence de polynômes de recalibration (cf. chapitre 5 pour une discussion détaillée). Enfin le code de minimisation appelé *SALTshaker* est le premier code d'entraînement public publié pour la communauté.

Les principales différences entre l'entraînement SALT2.4 (Betoule et al. 2014) et les modèles SALT3 récemment publiés proviennent du fait que les deux modèles sont entraînés sur des lots différents (K21 est plus important, et a été recalibré sur PanSTARRS).

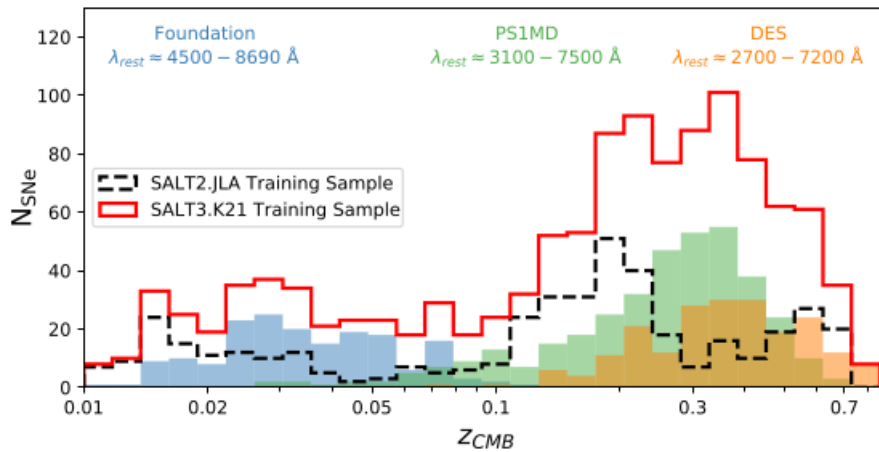


FIGURE 4.9 – Histogramme des *redshifts* de SNe Ia composant les lots d'entraînements de JLA et de l'ajout de K21 (figure extraite de Kenworthy et al. (2021)).

4.6 Conclusion

Les techniques de modélisation empirique des courbes de lumière et des spectres des SNe Ia ont considérablement évolué au cours des deux dernières décennies. Nous avons tenté d'en retracer succinctement l'évolution, depuis l'approche "historique" dans laquelle des modèles semi-analytiques étaient ajustés sur des données K -corrigées, jusqu'aux techniques modernes dans lesquelles les mesures sont directement prédites à l'aide d'un modèle complet, contenant une description de la diversité de la population des SNe Ia et une modélisation du processus de mesure conduisant aux spectres et aux flux large bande.

Depuis 2007, SALT2 constitue le modèle le plus complet utilisable dans le cadre d'une analyse cosmologique. Le modèle peut être entraîné simultanément sur des données spectroscopiques et photométriques. Il repose principalement sur des données photométriques, disponibles à tous les *redshifts*. Il ne fait par ailleurs aucune hypothèse sur la cosmologie sous-jacente. Cela permet de réaliser l'entraînement sur la totalité de la statistique du diagramme de Hubble, et non sur un sous-ensemble potentiellement non-représentatif, des SNe Ia utilisées pour l'analyse cosmologique (par exemple, les SNe Ia proches). D'autres modèles ont été développés, notamment ceux basés sur les séquences spectrophotométriques assemblées par le projet SNfactory. Ce type de données

n'étant malheureusement disponible que dans la gamme $z < 0.15$, il n'est pas possible d'entraîner les modèles SNEMO ou SUGAR sur la totalité de la statistique du diagramme de Hubble.

Avec l'avènement des projets Pantheon+ et surtout ZTF, la taille des lots disponibles pour entraîner les modèles a été multipliée par un facteur 5. À terme, LSST produira un lot homogène de plusieurs dizaines de milliers de supernovæ, avec un excellent suivi, dans la gamme de *redshifts* $0.05 < z < 1$. Ce mouvement vers les hautes statistiques est combiné à une augmentation significative de l'homogénéité et de la qualité des données de suivi.

L'augmentation de la statistique disponible impose de disposer à court terme d'un modèle utilisable dans la gamme de *redshifts* $z < 1$, pouvant être entraîné sur des lots de données de taille 10 à 20 fois supérieure à la taille des échantillons JLA ou K21. Par ailleurs, les études menées sur le lot SNfactory suggèrent fortement que la dimensionnalité de la population des SNe Ia est supérieure à 2.

Une voie assez naturelle est donc de revisiter la stratégie mise en place par SALT2 en conservant ce qui a fait le succès du modèle, à savoir un entraînement basé principalement sur des données photométriques, ce qui garantit son utilisabilité à tous les *redshifts*, en désignant le code de façon à rendre aisée l'implémentation de modèles de plus haute dimensionnalité et en accélérant et simplifiant autant que possible la procédure d'entraînement. Ce programme de travail a occupé l'essentiel de ma thèse. C'est l'objet des chapitres qui suivent.

Le modèle spectrophotométrique

La première étape de mon travail de thèse a été de développer un modèle spectrophotométrique hybride (i.e. pouvant être entraîné sur des spectres et des courbes de lumière). Nous sommes repartis de la paramétrisation du modèle SALT2 [Guy et al. (2007)] dans lequel la variabilité des SNe est paramétrée à l'aide d'un paramètre de forme et d'un paramètre de couleur. Le développement du modèle a été réalisé dans un langage de programmation moderne, `python3`¹, en utilisant massivement la bibliothèque de calcul vectoriel `numpy`² qui offre un accès aisé au parallélisme ainsi qu'à des méthodes d'algèbre linéaire rapides.

Nous décrivons tout d'abord les quantités physiques que nous cherchons à modéliser, et nous détaillons le processus de mesure, à savoir, le lien entre ces quantités physiques et les observables. Ensuite dans la section 5.2 nous présentons le modèle lui-même. Les incertitudes de mesure ne suffisent jamais à expliquer la dispersion des résidus d'entraînement : en d'autres termes, aucun modèle développé jusqu'ici ne permet de capturer toute la variabilité de SN à SN. Un second ingrédient de l'analyse est un modèle d'erreur, qui permet de décrire cette variabilité résiduelle, et de propager les incertitudes associées, jusqu'à l'ajustement des paramètres cosmologiques sur le diagramme de Hubble. La section 5.4 détaille le traitement des incertitudes de calibration. La section 5.5 décrit les astuces d'implémentation permettant de réduire le temps d'évaluation du modèle et de ses dérivées. Une brève section décrit enfin l'entraînement de ce modèle.

1. <https://www.python.org/download/releases/3.0/>

2. <https://numpy.org/>

Contents

5.1	De la luminosité de la SN aux données de suivi	119
5.2	Le modèle SALT2	120
5.2.1	Modèle spectroscopique	123
5.2.2	La modélisation photométrique	123
5.2.3	Les contraintes	124
5.2.4	Recalibration des spectres	125
5.2.5	La régularisation	127
5.3	Modèle d’erreur	127
5.3.1	<i>Error Snake</i>	128
5.3.2	Modélisation de la dispersion en couleur : <i>Color Scatter</i>	128
5.4	Propagation des incertitudes de calibration	129
5.5	Des équations au code	130
5.5.1	Implémentation des splines	131
5.5.2	Nombre de paramètres	132
5.5.3	Évaluation des flux photométriques	133
5.6	Entraînement du modèle	135
5.7	Conclusion	136

5.1 De la luminosité de la SN aux données de suivi

On cherche à modéliser la luminosité d'une supernova à un instant donné t_e et à une longueur d'onde donnée λ_e définis dans le référentiel de la supernova ($e = \text{émission}$). En pratique, on repère les temps par rapport au maximum de luminosité dans la bande B et on définit la phase de la SN comme :

$$p_e = t_e - t_{\max}^B \quad (5.1)$$

L'évolution de la luminosité de la supernova en fonction de la phase est décrite par la surface :

$$\mathcal{S}(\lambda_e, p_e) \quad (5.2)$$

\mathcal{S} est une quantité physique (*spectral flux*), homogène à une puissance par unité de longueur d'onde, exprimée par exemple en W/ ou en erg/s/. À partir de \mathcal{S} , d'une cosmologie et d'un modèle de notre instrumentation, il est possible de prédire les spectres et les courbes de lumière observés.

Spectres Lors de son voyage à travers l'Univers, un photon émis à une longueur d'onde λ_e subit le décalage vers le rouge dû à l'expansion cosmique. Dans le référentiel de l'observateur sa longueur d'onde est : $\lambda_o = (1+z)\lambda_e$ et son énergie est divisée par un facteur $1/(1+z)$. L'expansion de l'Univers induit également une dilatation des temps qui affecte non seulement les phases observées : $p_o = (1+z)p_e$ mais aussi les flux (le nombre de photons mesurés par unité de temps). Enfin, l'observateur observe un flux, à savoir un nombre de photons par unité de surface et de temps. La relation entre S_e et ce que mesure l'observateur est :

$$\phi_{\text{spec}}(\lambda_o, p_o) = \frac{1}{4\pi d_L^2(z)} \frac{1}{1+z} \mathcal{S}_e \left(\frac{\lambda_o}{1+z}, \frac{p_o}{1+z} \right) \quad (5.3)$$

où $d_L(z)$ est la distance de luminosité définie section 1.3.2.

ϕ_{spec} a les unités d'une *irradiance spectrale monochromatique*³ (Léna 1996) ou *spectral flux density*. Elle est homogène à une énergie par unité de temps, de surface et de longueur d'onde, par exemple : $\text{erg.s}^{-1}.\text{m}^{-2}.\text{\AA}^{-1}$.

Flux photométriques En pratique, les données utilisées pour mesurer des distances sont des données photométriques, intégrées dans des bandes d'une centaine de nanomètres de largeur. En notant $T_X(\lambda)$ la transmission de l'instrument dans la bande X , le flux mesuré sera :

$$\phi_{\text{phot}}(p_o) = \int S_o(\lambda_o, p_o) \frac{\lambda_o}{hc} T_X(\lambda_o) d\lambda_o \quad (5.4)$$

$$(5.5)$$

ou alors, en substituant l'expression de S_o donnée equation 5.3 et en changeant de variable pour intégrer dans le référentiel de la SN :

$$\phi_{\text{phot}}(p_o) = \frac{1}{4\pi d_L^2(z)} (1+z) \int \mathcal{S}_e \left(\lambda_e, \frac{p_o}{1+z} \right) \frac{\lambda_e}{hc} T_X((1+z)\lambda_e) d\lambda_e \quad (5.6)$$

3. L'irradiance spectrale monochromatique est appelée flux en astronomie, elle est souvent exprimée en Jansky [Js] égale à $10^{-26} \text{ W m}^{-2} \text{ Hz}^{-1}$.

Normalisation Les distances de luminosité des SNe sont généralement définies en fonction du flux émis dans les référentiels des SNe (*restframe*) et observé dans la bande B au maximum de luminosité, et plus exactement du flux *restframe* dans la bande B observé à 10pc de la SN :

$$\Phi_B = \frac{1}{4\pi(10\text{pc})^2} \int \mathcal{S}_e(\lambda, p = 0) \frac{\lambda}{hc} B(\lambda) d\lambda \quad (5.7)$$

En imposant la normalisation du modèle dans la bande B -*restframe*, aussi noté B^* , à phase 0 :

$$\int \mathcal{S}_e(\lambda, p = 0) \frac{\lambda}{hc} B(\lambda) d\lambda = 1 \quad (5.8)$$

et en définissant :

$$X_0 = \Phi_B \frac{(10\text{pc})^2}{d_L^2(z)} \quad (5.9)$$

on obtient :

$$\phi_{\text{spec}}(\lambda_o, p_o) = X_0 \frac{1}{1+z} \mathcal{S}_e \left(\frac{\lambda_o}{1+z}, \frac{p_o}{1+z} \right) \quad (5.10)$$

$$\phi_{\text{phot}}(p_o) = X_0 (1+z) \int \mathcal{S}_e \left(\lambda_e, \frac{p_o}{1+z} \right) \frac{\lambda_e}{hc} T_X((1+z)\lambda_e) d\lambda_e \quad (5.11)$$

Ce sont les équations de base de notre modèle spectrophotométrique. Il peut être utilisé pour prédire des données spectroscopiques ou photométriques. Le modèle ne présuppose aucune cosmologie (hormis l'effet du *redshift* sur les spectres et les flux). L'information cosmologique est entièrement encodée dans le paramètre X_0 , qui est ajusté sur les données de suivi de chaque SN.

5.2 Le modèle SALT2

L'élément central du modèle, est la surface spectrale $\mathcal{S}(\lambda, p)$ ⁴. La forme de cette surface varie légèrement d'une SN à l'autre. Mais nous savons que cette variabilité peut s'exprimer à l'aide de quelques paramètres seulement. Plusieurs modèles ont été proposés. Pour cette étude, nous avons choisi de repartir de la paramétrisation proposée par SALT2 :

$$S(\lambda, p) = [M_0(\lambda, p) + X_1 M_1(\lambda, p)] 10^{0.4cCL(\lambda)} \quad (5.12)$$

La paramétrisation proposée ci-dessus, est une décomposition classique de type "analyse en composantes principales", tronquée à l'ordre 1, et complétée par une loi de couleur indépendante de la phase.

M_0 est la surface spectrale moyenne (*restframe*) des supernovæ du lot d'entraînement (figure 5.1). La surface M_1 représente la principale variation à M_0 (figure 5.2). Ces deux surfaces sont communes à toutes les SN. Comme elles sont totalement empiriques, nous choisissons de les

4. Dans la suite nous omettons l'indice e .

décomposer sur une base de splines à deux dimensions. CL est une loi de couleur (figure 5.3), également commune à toutes les SNe. Elle est définie par un polynôme fonction de la longueur d'onde.

X_1 et c sont des paramètres propres à chaque SN, auxquels s'ajoute $t_{max}^{B^*}$ (caché dans la définition de la phase de la SN). $t_{max}^{B^*}$ est la date du maximum de luminosité en bande B^* , les temps étant mesurés dans le référentiel terrestre. X_1 est un paramètre de forme, bien corrélé au stretch, et c paramétrise des variations de couleur d'une SN à l'autre. La loi de couleur est définie de façon à ce que les paramètres c correspondent à une variation de couleur indépendante de la phase.

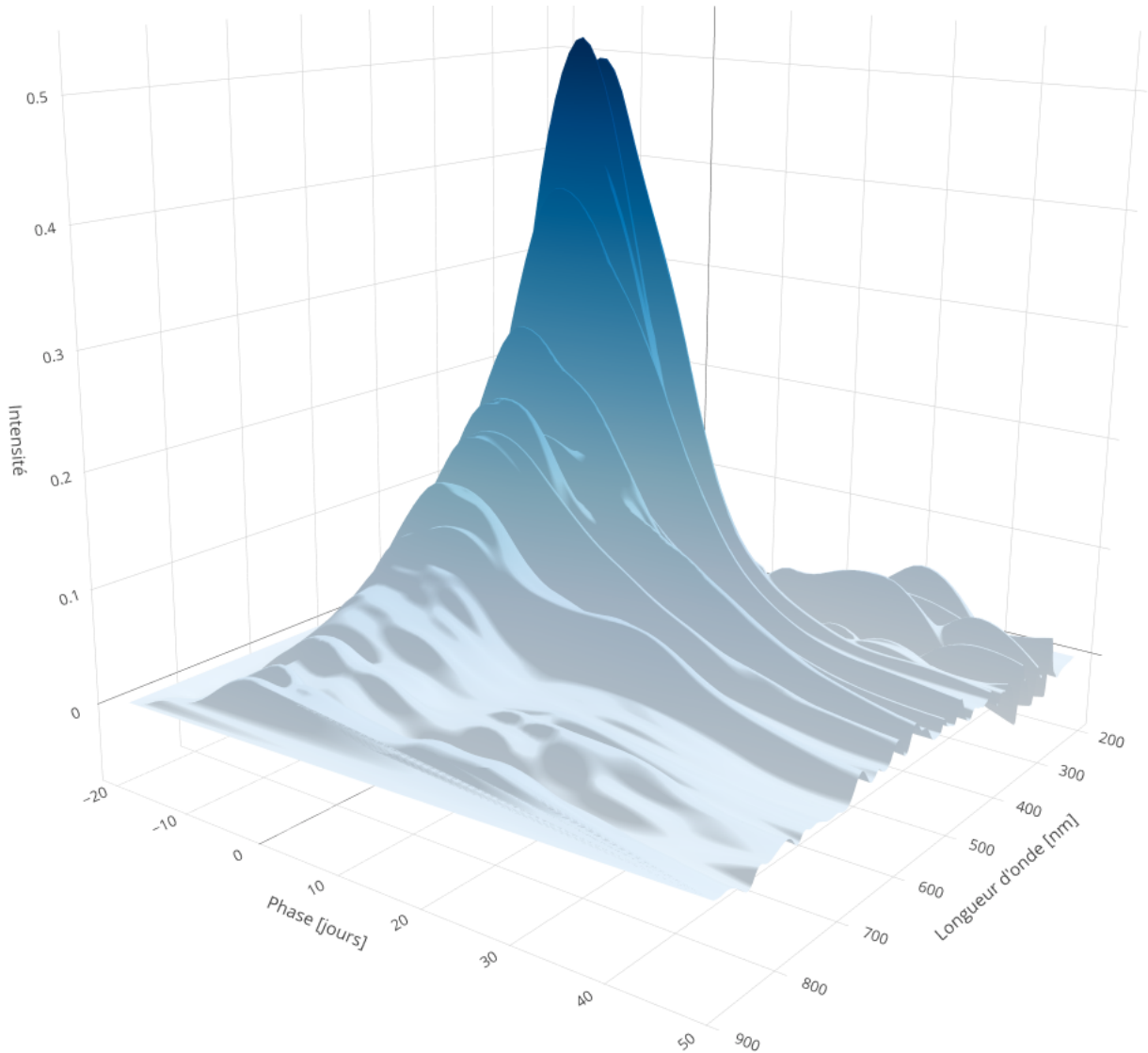


FIGURE 5.1 – Surface M_0 du modèle SALT2.4, en fonction de la phase et de la longueur d'onde.

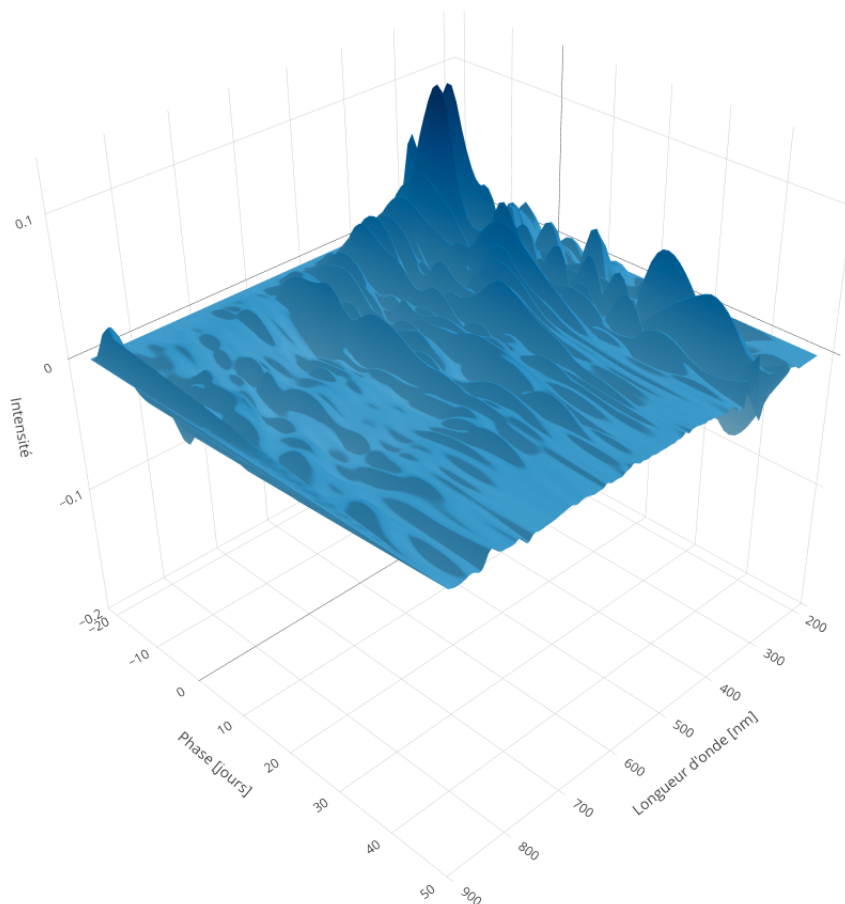


FIGURE 5.2 – Surface M_1 du modèle SALT2.4, en fonction de la phase et de la longueur d'onde.

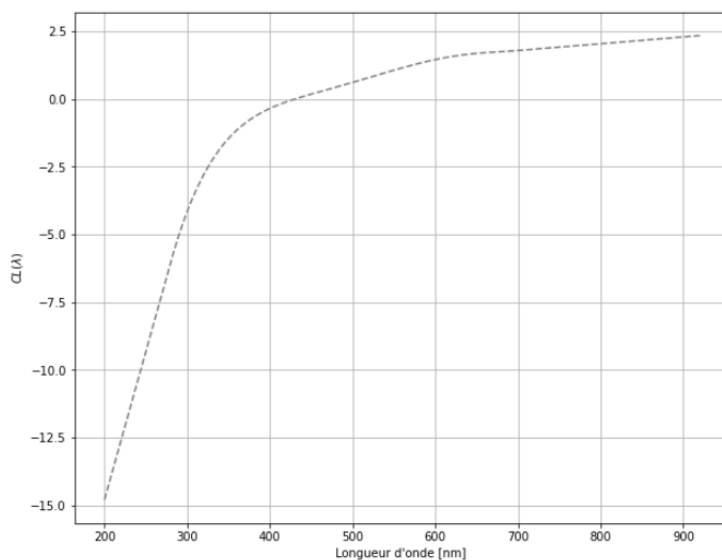


FIGURE 5.3 – Loi de couleur du modèle SALT2.4 en fonction de la longueur d'onde.

L'entraînement du modèle consiste à ajuster simultanément les paramètres globaux (M_0 , M_1 , CL) et les paramètres propres à chaque SN (X_0 , X_1 , c et $t_{\max}^{B^*}$) sur un jeu de données photomé-

triques et spectroscopiques.

5.2.1 Modèle spectroscopique

Le spectre d'une SN à une phase et un redshift z donnés s'écrit donc (équations 5.11 et 5.13) :

$$\phi_{\text{spec}}(\lambda, p) = \frac{X_0}{1+z} [M_0(\lambda, p) + X_1 M_1(\lambda, p)] 10^{0.4cCL(\lambda)} \quad (5.13)$$

La figure 5.4 montre le calcul d'un spectre observé à une phase égale à $p = 2.047$ jours et émis à $z = 0.0775$.

C'est une combinaison linéaire d'une coupe de M_0 et M_1 aux longueurs d'ondes et à la phase d'observation, corrigée de la loi de couleur évaluée à la longueur d'onde du spectre. Trois paramètres propres à la SN entrent en jeu, le paramètre de forme X_1 , la couleur c et le temps au maximum de luminosité en bande B^* . La colonne de gauche, montre la reconstruction du spectre divisé en trois étapes, le modèle moyen, corrigé de la variabilité de forme et ensuite de la loi de couleur. La colonne de droite correspond à l'évaluation des différentes grandeurs ajoutées au modèle moyen.

En pratique, les spectres sont entachés d'incertitudes de calibration en flux un ordre de grandeur plus élevées que les données photométriques (quelques pourcents, versus 0.1%). Au cours de l'entraînement, on choisit donc d'absorber l'information photométrique bande-large dans un polynôme de calibration, fonction de la longueur d'onde.

5.2.2 La modélisation photométrique.

De même, le flux d'une SN Ia à une phase p et un *redshift* z , observé par un instrument de transmission T_X , dans le référentiel de l'observateur s'écrit (équations 5.11 et 5.13) :

$$\phi_{\text{phot}}(p_o) = X_0(1+z) \int [M_0(\lambda, p) + X_1 M_1(\lambda, p)] 10^{0.4cCL(\lambda)} \frac{\lambda}{hc} T_X(\lambda(1+z)) d\lambda \quad (5.14)$$

Connaissant la transmission de l'instrument, T_X , (cf figure 3.10), le modèle peut prédire simplement le flux photométrique. La figure 5.5 illustre le calcul. Pour chaque point de photométrie, un spectre est généré (équation 5.13) à une phase donnée : ce sont les courbes en tirets noirs sur la figure 5.5. Le produit de ce spectre par la transmission de l'instrument (zone bleue pour le filtre g et rouge pour le filtre i) est ensuite intégré sur l'intervalle de longueur d'onde pour obtenir un point de photométrie, dont l'évolution forme la courbe de lumière (projetée sur l'arrière plan).

5.2.3 Les contraintes

Le modèle présenté ci-dessus est un modèle empirique. Les surfaces M_0 , M_1 et la loi de couleur CL possèdent un grand nombre de degrés de liberté de façon à pouvoir ajuster au mieux, et sans a priori, les données du lot d'entraînement. Cela induit plusieurs directions de dégénérescence entre paramètres globaux et paramètres propres à chaque SN. Celles-ci peuvent être supprimées en imposant un ensemble de contraintes bien choisies.

Contrainte sur l'amplitude du modèle Il ressort clairement des équations 5.13 et 5.14 que l'amplitude du modèle X_0 est dégénérée avec la normalisation des surfaces M_0 et M_1 , ainsi qu'avec la loi de couleur. En effet, en multipliant X_0 par une certaine quantité et en divisant M_0 et M_1 et CL par la même quantité, la valeur du modèle reste inchangée.

On choisit de normaliser le modèle, de façon à ce que son intégrale au pic de luminosité t_{\max}^B dans la bande B -restframe soit égale à 1 pour toutes les SNe. Cela implique :

$$\int M_0(\lambda, p = 0,) T_B(\lambda) \frac{\lambda}{hc} d\lambda = 1 \quad (5.15)$$

et

$$\int M_1(\lambda, p = 0) T_B(\lambda) \frac{\lambda}{hc} d\lambda = 0 \quad (5.16)$$

ainsi que

$$CL(\lambda^B) = 0 \quad (5.17)$$

Dans notre procédure d'entraînement, les contraintes sur M_0 et M_1 sont externes au modèle : elles sont implémentées sous forme de pénalités ajoutées au χ^2 de l'ajustement. Les contraintes sur la loi de couleur en revanche sont implémentées directement dans le modèle.

Contraintes sur la loi de couleur La loi de couleur elle-même est dégénérée avec le paramètre de couleur c , ainsi qu'avec les surfaces M_0 et (dans une moindre mesure) M_1 : il est possible de modifier simultanément les couleurs des SNe, la loi de couleur et les surfaces, sans changer la valeur du modèle. Pour briser cette dégénérescence, nous suivons la paramétrisation proposée par SALT2, en imposant la forme et la pente de $CL(\lambda)$ entre les bandes B et V .

La loi de couleur est définie sur l'intervalle $\lambda^{UV} = 2800 \text{ \AA} < \lambda < 7000 \text{ \AA} = \lambda^{IR}$ comme un polynôme de degré 5, tel que $CL(\lambda^B) = 0$ et $CL(\lambda^V) = 1$, en longueur d'onde définie dans le référentiel de la SN. En dehors de cet intervalle, étant donnée la faible statistique disponible, elle est prolongée par de simples lois affines de paramètres :

$$CL(\lambda) = \begin{cases} CL'(\lambda_r^{UV}) \times (\lambda_r - \lambda_r^{UV}) + CL(\lambda_r^{UV}) & \lambda < \lambda_{UV} \\ \lambda_r \times \left(\sum_{i=1}^4 p_i \lambda_r^i + 1 - \sum_{i=1}^4 p_i \right) & \lambda^{UV} < \lambda < \lambda^{IR} \\ CL'(\lambda^{IR}) \times (\lambda_r - \lambda_r^{IR}) + CL(\lambda_r^{IR}) & \lambda > \lambda^{IR} \end{cases} \quad (5.18)$$

avec

$$\lambda_r = \frac{\lambda - \lambda^B}{\lambda^V - \lambda^B} \quad (5.19)$$

Contrainte sur le domaine de phase D'autres dégénérescences existent entre les surfaces et les temps au maximum de luminosité en bande B, $t_{max}^{B^*}$, de chaque SN. Là encore, l'ajout d'une constante à tous les $t_{max}^{B^*}$ peut être compensé presque exactement par une modification des surfaces spectrales, qui laisserait le modèle inchangé. SALT2 (et SALT3) résolvent le problème en ajustant au préalable les dates du maximum de luminosité pour chaque SN, et les maintenant fixes durant l'entraînement. Nous avons choisi de conserver les dates de maximum dans l'ajustement, au prix d'une contrainte supplémentaire. On impose donc que $t_{max}^{B^*}$ soit un extremum de la courbe de lumière en bande B^* :

$$\left. \frac{\partial \phi_{\text{phot}, B^*}(t)}{\partial t} \right|_{t=t_{max}^{B^*}} = 0 \quad (5.20)$$

Ce qui se traduit par une contrainte directe sur les deux surfaces à $p = 0$:

$$\int \left. \frac{\partial M_0(\lambda, p=0)}{\partial t} \right|_{t=t_{max}^{B^*}} T_B(\lambda) \frac{\lambda}{hc} d\lambda = 0 \quad (5.21)$$

$$\int \left. \frac{\partial M_1(\lambda, p=0)}{\partial t} \right|_{t=t_{max}^{B^*}} T_B(\lambda) \frac{\lambda}{hc} d\lambda = 0 \quad (5.22)$$

Contrainte sur la distribution des X_1 Il reste encore à fixer la définition des paramètres X_1 , qui restent dégénérés avec l'échelle de la surface M_1 et avec X_0 et c . On impose donc que la moyenne des X_1 des SNe du lot d'entraînement soit nulle (i.e. la SN moyenne est de X_1 nul). On doit également fixer soit l'échelle des X_1 , soit celle de la surface M_1 . Nous avons choisi la première voie. Les deux contraintes suivantes sont donc ajoutées :

$$\langle X_1 \rangle = 0 \quad (5.23)$$

$$\langle (X_1 - \langle X_1 \rangle)^2 \rangle = 1 \quad (5.24)$$

Contrainte sur la distribution des c À ce stade, la dernière dégénérescence qui subsiste concerne les paramètres X_0 et c . Elle est brisée en imposant que la distribution des couleurs soit centrée :

$$\langle c \rangle = 0 \quad (5.25)$$

en d'autres termes, la supernova "moyenne" est de couleur 0. La couleur c caractérise un "rougissement" ou "bleuissement" relatif, par rapport à la supernova moyenne. Ainsi sur le lot d'entraînement, la moyenne de X_1 (respectivement des c) sera nulle par construction. La définition de la SN "moyenne", et donc du zéro des couleurs et des X_1 est une fonction de l'échantillon d'entraînement.

Le tableau 5.1 regroupe les contraintes explicites qui seront intégrées à l'étape de minimisation, sous forme de pénalisations quadratiques :

5.2.4 Recalibration des spectres

Comme indiqué plus haut, la calibration photométriques des spectres est généralement affectée de modes d'erreurs importants aux grandes échelles en longueur d'onde (cf section 3.2.3). Il est

Contraintes	dégénérescence entre :
$\int M_0(\lambda, p = 0,) T_B(\lambda) \frac{\lambda}{hc} d\lambda = 1$	X_0, X_1, c et $M_0(p = 0)$
$\int M_1(\lambda, p = 0,) T_B(\lambda) \frac{\lambda}{hc} d\lambda = 0$	X_0, X_1, c et $M_1(p = 0)$
$\int \left. \frac{\partial M_0(\lambda, p=0)}{\partial t} \right _{t=t_{max}^{B^*}} T_B(\lambda) \frac{\lambda}{hc} d\lambda = 0$	$t_{max}^{B^*}$ et M_0
$\int \left. \frac{\partial M_1(\lambda, p=0)}{\partial t} \right _{t=t_{max}^{B^*}} T_B(\lambda) \frac{\lambda}{hc} d\lambda = 0$	$t_{max}^{B^*}$ et M_1
$\langle c \rangle = 0$	X_0, X_1 et c
$\langle X_1 \rangle = 0$	X_0 et X_1
$\langle (X_1 - \langle X_1 \rangle)^2 \rangle = 1$	X_1 et M_1

TABLE 5.1 – Contraintes explicites sur les paramètres du modèle.

impératif de supprimer ces modes d'erreur, tout en préservant l'information apportée par les spectres aux petites échelles (*features spectrales*)

Ceci est réalisé lors de l'entraînement en multipliant les prédictions spectrales du modèle par un polynôme de recalibration spécifique à chaque spectre, fonction de la longueur d'onde d'observation λ_o , et d'ordre $N_s = 3$, commun à tous les spectres :

$$s(\lambda_{rec}) = \sum_i^{N_s} s_i \lambda_{rec}^{N_s - i} \quad (5.26)$$

avec :

$$\lambda_{rec} = \frac{2(\lambda_o - \lambda_{max})}{\lambda_{max} - \lambda_{min}} + 1 \quad (5.27)$$

Chaque polynôme est fonction de la longueur d'onde dans le référentiel de l'observateur.

Il est défini sur la plage spécifique $\lambda_{min} < \lambda < \lambda_{max}$ couverte par les données du spectre. Puisque tous les spectres sont recalibrés simultanément sur la photométrie, il n'est pas nécessaire d'imposer que le degré du polynôme soit égale au nombre de courbe de lumières de chaque SN. Il n'est pas, non plus nécessaire, d'avoir un point de photométrie à la date exacte de l'observation du spectre grâce au grand nombre de spectres utilisées dans l'entraînement. Dans cette configuration, nous nous assurons que lors de l'entraînement, les petites échelles des surfaces spectrales sont contraintes par les spectres, mais que les grandes échelles des mêmes surfaces ne sont contraintes que par la photométrie. Le prix à payer est $4 \times N_{spectra}$ paramètres supplémentaires, soit typiquement quelques milliers.

Du fait de cette recalibration, la présence du paramètre de normalisation X_0 dans la partie spectrale du modèle devient inutile, et on réécrit l'équation 5.13 sous la forme :

$$\phi_{spec,o}(\lambda_o, p_o) = \frac{1}{1+z} \left[M_0 \left(\frac{\lambda_o}{1+z}, \frac{p_o}{1+z} \right) + X_1 M_1 \left(\frac{\lambda_o}{1+z}, \frac{p_o}{1+z} \right) \right] 10^{0.4cCL(\lambda_r)} s(\lambda_{rec}) \quad (5.28)$$

5.2.5 La régularisation

Malgré la croissance soutenue du nombre de supernovae de qualité cosmologique, les lots d'entraînement ne permettent pas encore de peupler uniformément le plan phase-longueur d'onde (p, λ) sur lequel sont définies les surfaces spectrales.

Par ailleurs, nous avons délibérément choisi de définir le modèle sur une zone du plan (phase-longueur d'onde) plus importante que l'espace contenant des données. En effet, comme nous ajustons la date du pic de luminosité de chaque SN, la phase associée à chaque point de mesure est susceptible de changer au cours de la minimisation. La zone du plan (p, λ) contrainte évolue également au cours du ajustement. Afin d'éviter que des points entrent et sortent de l'espace de définition du modèle, nous étendons celui-ci au-delà de la zone contrainte par les données. Nous choisissons typiquement $2000 \text{ \AA} < \lambda < 9000 \text{ \AA}$ et $-20 < p < 50$.

Par conséquent, une fraction significative des coefficients des développements en splines des surfaces spectrales n'est pas contrainte par des données. Nous avons donc recours à une technique de régularisation. La régularisation injecte de l'information a priori, sur la régularité du modèle, afin de contraindre le modèle en l'absence de données.

En pratique, on ajoute au χ^2 une pénalité quadratique, portant sur la valeur du modèle (i.e. en l'absence de données, le modèle est censé tendre vers 0) et la valeur des dérivées du modèle (cf. section 6.2.2).

La procédure de régularisation permet de considérablement simplifier l'entraînement, en évitant d'adapter la base et le domaine de définition au lot de données d'entraînement, et en garantissant que toutes les données sont incluses dans l'espace de définition du modèle.

En revanche, la procédure de régularisation peut induire deux fragilités subtiles dans la procédure d'entraînement : tout d'abord, une régularisation mal réglée est susceptible de biaiser soit le modèle, soit les paramètres propres à chaque SN. Ce type de biais doit être évalué par exemple sur des échantillons simulés représentatifs du lot d'entraînement, comme nous le verrons dans les chapitres suivants. Ensuite, la régularisation peut masquer des quasi-dégénérescences du modèle. Une situation typique est celle où un spectre tardif, muni de son polynôme de recalibration est situé dans une zone où le modèle n'est tenu que par la régularisation.

5.3 Modèle d'erreur

Aucun des modèles empiriques existants ne décrit parfaitement les données d'entraînement. Plus précisément, à l'issue de l'entraînement, les incertitudes de mesure ne permettent pas d'expliquer la dispersion des résidus au modèle ajusté. Cela signifie qu'il existe une variabilité de SN à SN, portant sur la forme des courbes de lumière, ou sur la structure spectrale qui ne peut être capturée par un modèle à deux paramètres.

Il est donc indispensable de modéliser cette variabilité résiduelle intrinsèque des SNe par un modèle d'erreur. Celui-ci est une fonction empirique de la phase et de la longueur d'onde. Il dépend de paramètres qui doivent être ajustés au cours de la procédure d'entraînement.

Le modèle et le modèle d'erreur peuvent (et en principe doivent) être ajustés simultanément. L'ajustement des paramètres de la variance est algébriquement plus complexe que l'ajustement du modèle seul, mais reste faisable. Par ailleurs, c'est un problème non linéaire, qui conduit à un estimateur potentiellement biaisé, dont le biais doit, ici encore, être estimé sur des échantillons simulés.

Dans le cadre de cette étude, nous avons implémenté un modèle d’erreur simple, comme preuve de concept. Il comprend une partie purement diagonale (*error snake*), ainsi qu’une partie non diagonale, qui décrit un mode d’erreur spécifique (*color scatter*) observé sur les données d’entraînement.

5.3.1 Error Snake

La partie diagonale du modèle d’erreur (*error snake*) peut se comprendre comme la contribution des termes d’ordre supérieur de la PCA, ayant été tronqués. Elle est décrite comme une surface, définie dans le plan (λ, p) sur les mêmes intervalles de longueur d’onde et de phase que M_0 et M_1 .

La variance des points de données est composée du terme d’erreur sur la mesure du flux, σ_{err} , à laquelle on ajoute en quadrature une variance additionnelle $\sigma_{Mod}(\lambda, p)$:

$$V(\lambda, p) = \sigma_{Err}^2 + \sigma_{Mod}^2(\lambda, p) \quad (5.29)$$

avec :

$$\sigma_{Mod}^{spec} = g(\lambda, p) \times \phi_{spec}(p, \lambda) \quad (\text{spectres}) \quad (5.30)$$

$$\sigma_{Mod}^{phot} = g(\lambda, p) \times \phi_{phot}(p, \lambda) \quad (\text{photométrie}) \quad (5.31)$$

La surface $g(\lambda, p)$ est également développée sur une base de splines de dimension 2. Du fait de la statistique disponible, nous sommes contraints d’adopter pour $g(\lambda, p)$ une résolution plus faible que la résolution initiale de M_0 et M_1 .

Les paramètres de ce modèle d’erreur sont ajustées simultanément avec le modèle. Le chapitre suivant précise les équations normales associées à cet ajustement.

5.3.2 Modélisation de la dispersion en couleur : Color Scatter

Il a également été observé [Guy et al. (2007), Chotard et al. (2011), Kenworthy et al. (2021)] qu’un seul paramètre de couleur ne permet pas de décrire la totalité de la variabilité en couleur des supernovæ de type Ia. Il existe une variabilité résiduelle autour de la loi de couleur. Cette variabilité corrèle tous les points d’une même courbe de lumière, dans une bande donnée. Son amplitude dépend de la longueur d’onde *restframe* correspondant à la bande d’observation pour une SN de *redshift* donné.

Il existe plusieurs façons de prendre en compte cette variance additionnelle dans le processus d’entraînement. Une possibilité est de densifier la matrice de covariance des données, en y ajoutant des paramètres non diagonaux, corrélant les points de photométries d’une même bande. Une telle solution est coûteuse en temps de calcul : la matrice de covariance étant non diagonale, le calcul de son inverse nécessite au minimum une factorisation de Cholesky supplémentaire à chaque itération. Le calcul des dérivées par rapport aux paramètres du modèle d’erreur est également plus coûteux (cf. chapitre suivant).

Une méthode algébriquement équivalente, consiste à modéliser les corrélations observées dans les résidus à l’aide de paramètres libres ajoutés au modèle lui-même, tout en contraignant la variabilité de ces paramètres à l’aide d’un a priori, qui est lui-même ajusté durant l’entraînement. C’est cette seconde méthode a été choisie par souci de rapidité et de facilité d’implémentation.

Pour modéliser la variabilité résiduelle de la couleur des SNe non décrite par le modèle, on autorise donc l'amplitude relative de chaque courbe de lumières à varier par une quantité $(1 + \kappa)$:

$$\phi_{\text{phot}}(p_0) = X_0(1 + z)(1 + \kappa) \int S(\lambda, p) T(\lambda) \frac{\lambda}{hc} d\lambda \quad (5.32)$$

où κ est un paramètre additionnel, dépendant de la SN et de la bande d'observation.

La variabilité des κ est définie par un a priori gaussien centré et dont la variance dépend de la longueur d'onde *restframe*, et doit être déterminée lors de l'ajustement. Dans notre modèle, ce terme est implémenté par un polynôme de la longueur d'onde *restframe* de la SN. Pour une courbe de lumière observée dans un bande X , de longueur d'onde moyenne λ^X , la variance correspondante est :

$$\sigma_{\kappa}^2 = P\left(\frac{\lambda^X}{1 + z}\right) \quad (5.33)$$

où z est de *redshift* de la supernova et P est un polynôme de la longueur d'onde *restframe*. En pratique, P est implémenté en fonction d'une longueur d'onde *restframe* réduite : λ_r définie par l'équation 5.19.

On présente figure 5.6, la dispersion en couleur en fonction de la longueur d'onde pour les entraînement de SALT2 et SALT3. Le modèle SALT2 paramétrise cette dispersion par un polynôme en longueur d'onde de degré 2 et SALT3 par un polynôme de degré 3 (cf figure 5.6). Cette différence s'explique par les différences entre lots d'entraînement, guidant la paramétrisation du modèle d'erreur.

5.4 Propagation des incertitudes de calibration

Toutes les mesures de flux sont affectées par les incertitudes portant sur la calibration photométrique des instruments utilisés pour le suivi. L'origine de ces incertitudes a été explicitée au chapitre 3. Elles sont liées à la chaîne de métrologie entre standards de flux primaires, calibrateurs secondaires (étoiles de champ) et courbes de lumière, ainsi qu'à la modélisation des spectres des standards de flux primaires.

Les chaînes de métrologie des flux mises en place par les relevés modernes sont précises à quelques pour mille (0.3% typiquement). La question de la précision de la modélisation des spectres de calibrateurs primaires est plus complexe. Les incertitudes qui impactent la cosmologie sont celles qui portent sur la pente des spectres modélisés, dans la gamme $g - z$. Les mises à jour récentes de la librairie CALSPEC conduisent à évaluer cette incertitude à $\sim 0.5\%$ [Brout et al. (2022)], soit un facteur 3 au-dessus du bilan d'erreur officiel de CALSPEC [Bohlin (2014)].

Les incertitudes liées à la chaîne de métrologie corrént toutes les mesures photométriques prises dans une même bande-passante. Les incertitudes portant sur la pente des standards de flux primaires corrént toutes les mesures photométriques prises dans des bandes différentes. La matrice de covariance de la contribution des erreurs de calibration aux mesures individuelles est donc dense. En revanche son rang est faible : il est égal au nombre de bandes N_b présentes dans le lot d'entraînement. Le bilan total des incertitudes de calibration peut-être compressé dans une matrice $N_b \times N_b$ que nous appellerons V_{cal} .

Une erreur de calibration photométrique est susceptible de distordre la surfaces M_0 ainsi que la loi de couleur. Elle génèrera des modes d'erreur complexes affectant les paramètres X_0 et C de toutes les SNe dont les courbes de lumière seront ajustées avec ce modèle. Il est donc essentiel de propager, lors de l'entraînement, le bilan des incertitudes de calibration jusqu'au modèle global ainsi qu'aux paramètres propres à chaque SN.

La méthode classique, implémentée par exemple dans [Guy et al. (2010); Betoule et al. (2014)], consiste à calculer les dérivées numériques du modèle en fonction des paramètres de calibration, puis utiliser ces dérivées pour propager les incertitudes de calibration aux paramètres propres à chaque SN. Cette méthode est simple conceptuellement, mais coûteuse en temps de calcul : si le lot d'entraînement contient 20 bandes, 20 ré-entraînements sont nécessaires pour produire la matrice des dérivées du modèle.

Une méthode alternative consiste à modéliser l'impact des incertitudes de calibration directement dans le modèle, à l'aide de paramètres spécifiques, dont les variations sont contraintes par un prior, dont la variance est donnée par le bilan d'incertitudes du processus de calibration. La méthode est extrêmement semblable à la modélisation du *color scatter* à ceci près que cette fois-ci le prior est connu : c'est un résultat de la calibration du *survey*.

L'équation du modèle photométrique devient :

$$\phi_{\text{phot}}(p_0) = X_0(1+z)(1+\kappa)(1+\eta) \int S(\lambda, p) T(\lambda) \frac{\lambda}{hc} d\lambda \quad (5.34)$$

où le paramètre η décrit l'effet d'une erreur de calibration dans la bande considérée. Au total, nous devons ajouter au modèle un paramètre η par bande. Ces nouveaux paramètres seront tenus dans le χ^2 total par un prior gaussien, dont l'algèbre sera détaillé au chapitre suivant.

Il existe des incertitudes de calibration d'un second type : celles qui portent sur la détermination des bandes passantes du survey. Elles impactent l'analyse en deux endroits distincts : les prédictions du modèle lui même, et la détermination de l'échelle de flux dans une bande donnée à partir des spectres des standards primaires. Il été montré [Hazenbergh (2019)] que la précision de la détermination des bandes passantes des surveys a un impact significatif sur les incertitudes finales. Pour un diagramme de Hubble contenant de l'ordre de 10^4 supernovæ, la précision requise sur le positionnement en longueur d'onde des bandes passantes est de l'ordre de 1\AA . Pour des raisons de temps, nous avons choisi de ne pas inclure la propagation de ces incertitudes dans cette étude. Il est possible de le faire, en suivant un schéma analogue.

5.5 Des équations au code

Le code est écrit en `python3` et utilise intensivement les bibliothèques de calcul numérique `numpy` and `scipy` qui permettent d'atteindre des performances en calcul pur proches de celles du langage C tout en permettant une souplesse algorithmique largement supérieure. À toutes les étapes du développement, nous avons essayé de minimiser autant que possible le temps d'exécution du modèle et de ses dérivées, afin de rendre l'entraînement aussi rapide que possible.

Dans cette section, nous détaillons quelques aspects pratiques qui nous ont permis d'accélérer l'exécution du code, principalement via quelques approximations, et via le précalcul et le stockage de quantités-clef souvent réutilisées.

5.5.1 Implémentation des splines

Les surfaces du modèle et du modèle d'erreur sont des fonctions totalement empiriques. Elles sont développées sur une base de splines, ce qui permet de préserver un nombre de degrés de liberté important (fixé par la densité de la base), tout en assurant leur continuité et la dérivabilité.

Les splines sont des fonctions définies par morceaux par des polynômes. Leur souplesse leur permet d'interpoler des formes complexes. Une spline d'ordre N est une fonction définie par morceaux par des polynômes en x , de degré $N - 1$, le recollement entre les morceaux étant C^{N-1} .

Les B-splines forment une famille particulière de splines : étant donnée une grille, on peut définir des splines d'ordre 1, positives à support compact par :

$$\mathcal{B}_{i \in [1, N], 1}(x) = \begin{cases} 1 & \text{si } n_i \leq x < n_{i+1} \\ 0 & \text{sinon.} \end{cases} \quad (5.35)$$

Puis, par récurrence à tous les ordres $m > 1$:

$$\mathcal{B}_{i \in [1, N], m+1}(x) = w_{i,m}(x)\mathcal{B}_{i,m}(x) + (1 - w_{i+1,m}(x))\mathcal{B}_{i+1,m}(x) \quad (5.36)$$

avec :

$$w_{i,k}(x) = \begin{cases} \frac{x-n_i}{n_{i+k}-n_i} & \text{si } n_{i+k} \neq n_i \\ 0 & \text{sinon.} \end{cases} \quad (5.37)$$

L'espace engendré par les \mathcal{B}_i d'ordre m forme un espace vectoriel dont les \mathcal{B}_i forment une base. On parle de B-splines. Le choix de la grille initiale détermine la résolution de la base. Le choix de l'ordre détermine la classe de régularité et l'existence et la continuité des dérivées de la fonction.

On peut voir sur la figure 5.7, les splines d'ordre respectif 1, 2 et 3 sur une grille de 5 nœuds et 11 nœuds. On remarque que les splines d'ordre 1 sont des fonctions-porte qui ne se recouvrent pas. Pour les surfaces M_0 et M_1 , nous utilisons une base de B-splines à 2 dimensions définies par le produit dyadique d'une base de splines définie sur l'intervalle de longueur d'onde, noté \mathcal{B}^m et une autre définie sur l'intervalle de phase, noté \mathcal{B}^t .

$$M_{0|1}(\lambda, p) = \sum_{k\ell} \theta_{k\ell}^{0|1} \mathcal{B}_k^t(p) \mathcal{B}_\ell^m(\lambda) \quad (5.38)$$

Les B-splines sont utilisées pour définir des formes complexes à l'aide de peu de paramètres, ou lisser des données bruitées. Elles nous permettent d'implémenter un modèle complètement empirique, sans aucun autre a priori que celui du choix de la grille, qui peut être ajusté sur des données bruitées via une méthode de moindres carrés.

Une propriété intéressante des B-splines qui apparaît lors d'un tel ajustement, est que la matrice Jacobienne du modèle (matrice des dérivées par rapport aux paramètres) est extrêmement peu dense, on dit qu'elle est creuse (*sparse*). Plus précisément, considérons une base $(\mathcal{B}_i(x))_{1 \leq i \leq n}$ de dimension n , un ensemble de points $(x_i)_{1 \leq i \leq N}$ de taille $N \gg n$ et la fonction $f(x) = \sum_{j=n}^N \theta_j(x)$ évaluée sur les x_i . La matrice du gradient de f évalué sur les x_i :

$$J_{ij} = (\mathcal{B}_j(x_i)) \quad (5.39)$$

est généralement extrêmement creuse. Cela provient du fait que les \mathcal{B}_i sont à support compact. Il en résulte que la Hessienne

$$H = J^T C^{-1} J \quad (5.40)$$

où C est la matrice de covariance des incertitudes sur les données est aussi creuse. Cette propriété permet d'accélérer considérablement l'inversion de la Hessienne qui est l'étape clef d'un ajustement par moindres carrés.

Il existe de nombreuses implémentations des bases de B-splines. Malheureusement, aucune de celles que nous avons testées ne permet d'évaluer et de retourner la matrice jacobienne sous forme d'une matrice creuse. Nous avons donc ré-implémenté la gestion et l'évaluation des bases de splines 1D et 2D sous forme d'une série de classes `python`. Afin d'accélérer l'évaluation, l'algorithme de deBoor est implémenté en `C`. À l'ordre 3, en dimension 1, la rapidité de l'évaluation est proche de 100 ns par point, ce qui place cette implémentation à peu près au niveau du `C`. En revanche, nous n'exploitons pas encore le parallélisme, ce qui permettrait d'accélérer encore l'évaluation du modèle.

5.5.2 Nombre de paramètres

Modèle Les surfaces M_0 (figure 5.1) et M_1 (figure 5.2) sont développées sur une base de splines d'ordre 4 à deux dimensions, commune aux deux surfaces, noté $\mathcal{B}_i(\lambda, p)$:

$$M_{0|1} = \sum_i \theta_{i,0|1} \mathcal{B}_i(\lambda, p) \quad (5.41)$$

où θ_0 et θ_1 sont les coordonnées de M_0 et M_1 sur la base. Ces paramètres sont communs à toutes les SNe. La base de spline en phase est définie par 20 nœuds entre -20 et $+50$ jours. La base en longueur d'onde est de 127 nœuds entre 2000 et 9200 \AA . Chaque surface est définie par 2838 paramètres.

La loi de couleur (5.3) est un polynôme d'ordre 5, sur l'intervalle de définition [$\lambda^{UV} = 2800 \text{ \AA}$, $\lambda^{IR} = 7000 \text{ \AA}$], prolongée affinement au-delà, et valant 0 en $\lambda = \lambda^B$ et 1 en $\lambda = \lambda^V$. La loi de couleur est définie par 4 paramètres (cf. équation 5.18).

La recalibration des spectres introduit 4 paramètres pas spectres

Modèle d'erreur La surface $g(p, \lambda)$ est développée sur une base de splines, d'ordre 1, que nous noterons $\mathcal{G}(\lambda, p)$:

$$g(\lambda, p) = \sum_i \gamma_i \mathcal{G}^i(\lambda, p) \quad (5.42)$$

où les paramètres γ_i sont les coordonnées de g sur la base \mathcal{G}_i .

La base de splines de l'*error snake* est définie en phase par 5 nœuds entre -20 et $+50$ jours et en longueur d'onde par 5 nœuds entre 2000 et 9200 \AA , soit 25 paramètres. La résolution de cette surface est bien plus faible que celle des surface M_0 et M_1 . Elle pourra augmenter avec la taille des lots d'entraînement.

La dispersion des paramètres κ est un polynôme de degré 3, de la longueur d'onde centrale de la bande définie dans le référentiel de la SN. Cela ajoute 3 paramètres :

$$\sigma^\kappa(\lambda_r) = \sum_{i=0}^{N=3} \sigma_i^\kappa \lambda_r^i \quad (5.43)$$

Si on considère un lot d'entraînement de N_{SN} supernovæ avec au total $N_{spectres}$ spectres, et N_{lcs} courbes de lumière observées dans N_{bandes} instruments, le nombre de paramètres total est donné par le tableau suivant :

Paramètres	# de paramètres		
modèle	NaCl	SALT2.4	SALT3
$\theta_0(N_{phase}, N_\lambda)$	2838 (22 × 129)	1287 (13 × 99)	2838 (22 × 129)
$\theta_1(N_{phase}, N_\lambda)$	2838 (22 × 129)	1287 (13 × 99)	2838 (22 × 129)
p	4	4	4
X_0	N_{SN}	N_{SN}	N_{SN}
X_1	N_{SN}	N_{SN}	N_{SN}
c	N_{SN}	N_{SN}	N_{SN}
$t_{max}^{B^*}$	N_{SN}	-	-
s_i	$4 \times N_{spectres}$	$\propto N_{spectres}$	$\propto N_{spectres}$
η	N_{bandes}	-	-
κ	N_{lcs}	-	-
Modèles d'erreur			
γ	25	60	(72 + 72 + 72)
σ_κ	4	4	4

TABLE 5.2 – Nombre de paramètres du modèle en fonction du nombre de SNe : N_{SN} , du nombre de courbes de lumière : N_{lcs} , du nombre de spectres : $N_{spectres}$ et du nombre de bandes : N_{bandes} .

La figure 5.8 montre l'évolution du nombre de paramètres en fonction du nombre de SNe et du nombre de spectres par SN. Pour le lot considéré, chaque SN a quatre courbes de lumière et le nombre d'instruments est fixé à 12. L'étoile labellée "JLA", correspond aux données du lot d'entraînement de SALT2.4 : 282 spectres et 1994 courbes de lumières, observées dans 30 instruments, pour 426 SNe ce qui représente 11361 paramètres au pour le modèle. L'étoile "K21", correspond au lot d'entraînement de SALT3, composé de 1285 spectres et 4695 courbes de lumières de 1083 SNe et donc 19960 paramètres à contraindre.

5.5.3 Évaluation des flux photométriques

Pour l'évaluation des flux photométriques, nous avons recours à une astuce qui permet d'accélérer significativement le calcul des intégrales. Le but est de dissocier les sections du calcul qui dépendent des paramètres du modèle, de celles qui n'en dépendent pas et qui peuvent donc être pré-calculées et stockées. D'après la formule 5.34, nous cherchons à calculer :

$$X_0 \times (1 + z) \times (1 + \kappa) \times (1 + \eta) \times [I_0 + X_1 I_1] \quad (5.44)$$

où l'on a :

$$I_{0|1} = \int M_{0|1}(\lambda, p) \frac{\lambda}{hc} T((1+z)\lambda) 10^{0.4 c CL(\lambda)} d\lambda \quad (5.45)$$

En pratique, on peut décomposer la transmission des filtres sur une base de splines : $T(\lambda) = \sum_q t_q \mathcal{B}_q(\lambda)$. En remplaçant cette expression dans l'équation 5.45, en sortant la partie temporelle de l'intégrale et en utilisant l'équation 5.38, on obtient :

$$I = \sum_{k\ell q} \theta_{k\ell} t_q \mathcal{B}_k(p) \int \mathcal{B}_\ell(\lambda) \frac{\lambda}{hc} \mathcal{B}_q((1+z)\lambda) 10^{0.4 c CL(\lambda)} d\lambda \quad (5.46)$$

Comme la loi de couleur est une fonction lentement variable de la longueur d'onde (comparée aux fonctions de base), nous pouvons la considérer constante sur l'espace d'intégration :

$$I \approx \sum_{k\ell q} \theta_{k\ell} t_q \mathcal{B}_k(p) \int \mathcal{B}_\ell(\lambda) \frac{\lambda}{hc} \mathcal{B}_q((1+z)\lambda) d\lambda \times 10^{0.4 c CL(\bar{\lambda})} \quad (5.47)$$

où $\bar{\lambda}$ une valeur judicieusement choisie de λ sur l'intersection des supports de $\mathcal{B}_\ell(\lambda)$ et $\mathcal{B}_q((1+z)\lambda)$. En fait, on peut montrer que si l'on choisit :

$$\bar{\lambda} \equiv \lambda_{\ell q}^{\text{eff}} \equiv \frac{\int \lambda^2 \mathcal{B}_\ell^m(\lambda) \mathcal{B}_q(\lambda(1+z)) d\lambda}{\int \lambda \mathcal{B}_\ell^m(\lambda) \mathcal{B}_q((1+z)\lambda) d\lambda} \quad (5.48)$$

l'approximation à l'ordre 2 est amplement suffisante, étant donnée la taille des supports des fonctions de base comparée aux variations de la loi de couleur.

L'intégrale ci-dessus se ré-écrit donc :

$$I \approx \sum_{k\ell q} \theta_{k\ell} t_q \mathcal{B}_k(p) G_{\ell q}[z] \times 10^{0.4 c CL(\lambda_{\ell q}^{\text{eff}})} \quad (5.49)$$

La matrice $G_{\ell q}[z]$ est appelée le Grammien des bases \mathcal{B}_ℓ et \mathcal{B}_q utilisées pour développer la partie spectrale du modèle et les filtres respectivement. Elle est extrêmement creuse étant donné que les fonctions des deux bases sont à support compact. Chaque intégrale est pré-calculée par quadrature de Gauss et stockée. Comme nous intégrons des polynômes, nous avons la garantie que l'utilisation d'une quadrature de Gauss d'ordre égale au degré des polynômes nous donne l'intégrale exacte. L'évaluation des intégrales $I_{0|1}$ se réduit donc à une contraction des paramètres avec des quantités pré-calculées.

Dans la version présentée ci-dessus, un Grammien distinct doit être calculé et stocké pour chaque SN. Une approche plus économique consiste à décomposer non pas $T(\lambda)$, mais $T(\lambda(1+z))$ sur la base de splines \mathcal{B}_q . Pour chaque SN ne sont stockés que les paramètres de la décomposition, $t_q[z]$, tandis que le Grammien, ainsi que la grille $\lambda_{\ell q}^{\text{eff}}$ restent universels.

$$I \approx \sum_{k\ell q} \theta_{k\ell} t_q[z] \mathcal{B}_k(p) G_{\ell q} \times 10^{0.4 c CL(\lambda_{\ell q}^{\text{eff}})} \quad (5.50)$$

C'est cette approche qui est implémentée dans la version actuelle du modèle. La base choisie comprend 900 nœuds entre 2000 et 11000 Å, soit un nœud tous les 10 Å. Nous avons vérifié en faisant varier la résolution de la base de 1 à 100 Å, que la différence de magnitude modélisée pour différent intervalle de *redshift* et pour différentes bandes ne dépasse pas 0.5*mmag* (sauf pour certaines bande en *U*, dont la limite est 1*mmag*). Pour des questions de rapidité, nous avons choisi une base de 900 nœuds sans aucun compromis sur la précision attendu.

La figure 5.9 montre que le temps d'évaluation du modèle et de ses dérivées croît avec le nombre de courbes de lumière car elles sont évaluées une par une (cf équation 5.50). En revanche, le

modèle spectral est évalué en une seule fois. Dans la plage statistique où nous avons testé le modèle, le temps d'évaluation ne semble pas croître linéairement avec le nombre de spectres. Ces temps d'évaluation ont été obtenues en utilisant un CPU de modèle : IntelCore™ i7-8665U à 1.90GHz.

5.6 Entraînement du modèle

Le modèle décrit dans cette section doit être ajusté sur des données dites *d'entraînement*, photométriques et spectroscopiques. Le modèle a été conçu de façon à ce que tous les paramètres puissent être contraints simultanément lors de cet ajustement. Pour mémoire, nous avons cinq grandes classes de paramètres :

1. les paramètres globaux du modèle proprement dit : (M_0, M_1, CL)
2. les paramètres propres à chaque SN $(X_0, X_1, c$ et $t_{\max}^{B^*})$
3. les paramètres du modèle d'erreur $(\gamma_i, \sigma_\kappa)$
4. les variables latentes κ et η
5. et enfin, les paramètres des polynômes de recalibration des spectres (s_i) .

Entraînement traditionnel La phase d'entraînement au sens traditionnel du terme, consiste à déterminer les paramètres globaux (M_0, M_1, CL) ainsi que les paramètres du modèle d'erreur $(\gamma_i, \sigma_\kappa)$, en marginalisant sur les autres classes de paramètres. L'entraînement est généralement organisé en plusieurs étapes distinctes :

1. détermination de la date du pic $(t_{\max}^{B^*})$ de chaque SN,
2. ajustement du modèle (sans modèle d'erreur)
3. analyse des résidus d'entraînement et détermination d'un premier modèle d'erreur
4. ajustement du modèle avec modèle d'erreur
5. re-détermination du modèle d'erreur
6. ré-itérer à partir de (4)

Le modèle ainsi déterminé est mis à la disposition de la communauté, et peut être utilisé pour l'ajustement des courbes de lumière des SNe et la détermination des paramètres propres à chaque SN, puis des distances aux SNe.

La technique a l'avantage de séparer clairement la phase d'entraînement, dont les itérations complexes requièrent un certain savoir-faire, de l'utilisation du modèle pour ajuster les courbes de lumière de chaque SN. Elle permet de disséminer dans la communauté un outil simple à utiliser. En revanche, la propagation des incertitudes est plus complexe : les incertitudes statistiques affectant le modèle et la dispersion en couleur sont propagées après le calcul des covariances entre les paramètres du modèle et les paramètres de l'ajustement des courbes de lumière de chaque SN (cf. Appendix A.3 [Guy et al. \(2010\)](#)). La propagation des incertitudes de calibration requiert de nombreux entraînements annexes afin de construire la matrice des dérivées nécessaire au calcul, comme discuté section 5.4.

L'entraînement proposé On notera que, *in fine*, les seuls paramètres d'intérêt sont les paramètres X_0, X_1 et c , utilisés pour construire des distances de luminosité standardisées ; et que, au cours de l'analyse, nous sommes amenés à marginaliser sur tous les autres paramètres du problème, sans exception. Le modèle lui-même, n'est que l'ingrédient qui stipule que l'évolution spectrophotométrique des SNe Ia est la même pour tous les objets, à deux degrés de liberté près.

Si le code d'entraînement est suffisamment rapide, et la procédure d'entraînement suffisamment simple, il est donc extrêmement tentant de cesser de dissocier l'étape d'entraînement de l'étape d'ajustement des courbes de lumière. Cela permet de s'assurer que toute l'information disponible est bien utilisée pour contraindre le modèle. Et surtout, cela permet de simplifier considérablement la gestion des incertitudes : celles-ci sont propagées de manière exacte, à partir du moment où nous sommes capables de construire et inverser la matrice Hessienne du problème.

Notre but est d'implémenter exactement cela : un code d'entraînement rapide, qui prenne en entrée un ensemble de courbes de lumière et de spectres assez consistant pour contraindre l'ensemble du modèle, et qui fournisse le vecteur des paramètres X_0, X_1, c de toutes les SNe, accompagné de sa matrice de covariance (après marginalisation sur tous les autres paramètres).

À noter que le lot d'entraînement ne se réduit pas nécessairement aux supernovæ présente *in fine* dans le diagramme de Hubble. Il peut contenir une statistique plus ample, en particulier des supernovæ extrêmement bien mesurées (e.g. 2011fe), mais trop proches pour appartenir au flot de Hubble.

5.7 Conclusion

Nous avons présenté l'implémentation d'un modèle empirique de l'évolution spectrophotométrique des SNe Ia. La paramétrisation que nous avons choisie est extrêmement proche de celle des modèles SALT2 et SALT3. Elle est évidemment appelée à évoluer, en fonction de ce que révéleront les entraînements sur les grands lots de données en cours de construction, notamment l'échantillon ZTF. Lors du développement du code, nous avons veillé à ce que les interfaces entre le modèle et le code de minimisation soient bien définies, afin de rendre simple le développement de nouvelles paramétrisations.

Le modèle est conçu de manière à ce que la totalité de l'entraînement puisse être réduite à une unique minimisation. En particulier, l'ajustement de la date du pic de luminosité des SNe n'est plus une étape séparée. Le code a également été écrit de façon à ce que le modèle d'erreur puisse être ajusté au cours de la même minimisation. Évidemment, une question ouverte à ce stade est de savoir si en pratique, les données permettent de contraindre les deux simultanément.

L'entraînement que nous cherchons à réaliser consiste en l'ajustement sous contraintes d'un modèle non-linéaire, assorti d'un modèle d'erreur dont les paramètres figurent dans la matrice de covariance des mesures. Nous souhaitons disposer d'un minimiseur rapide, capable de gérer des problèmes de taille 10^4 à 10^5 paramètres. Nous souhaitons qu'il puisse réaliser l'ajustement d'un modèle non linéaire sous contraintes, qu'il puisse gérer la régularisation (optionnelle) d'une partie des paramètres, et qu'il implémente l'algèbre spécifique à l'ajustement des variances. Enfin, étant donné la dimensionnalité importante du modèle, le minimiseur doit pouvoir tirer parti du caractère creux de la Hessienne du problème. Une recherche bibliographique menée fin 2019-début 2020 ne nous a pas permis de trouver une bibliothèque capable de réaliser cela. Nous avons donc écrit notre propre cadre de minimisation. C'est l'objet du prochain chapitre.

Spectre observé à $p = -2.047$ émit à $z = 0.775$

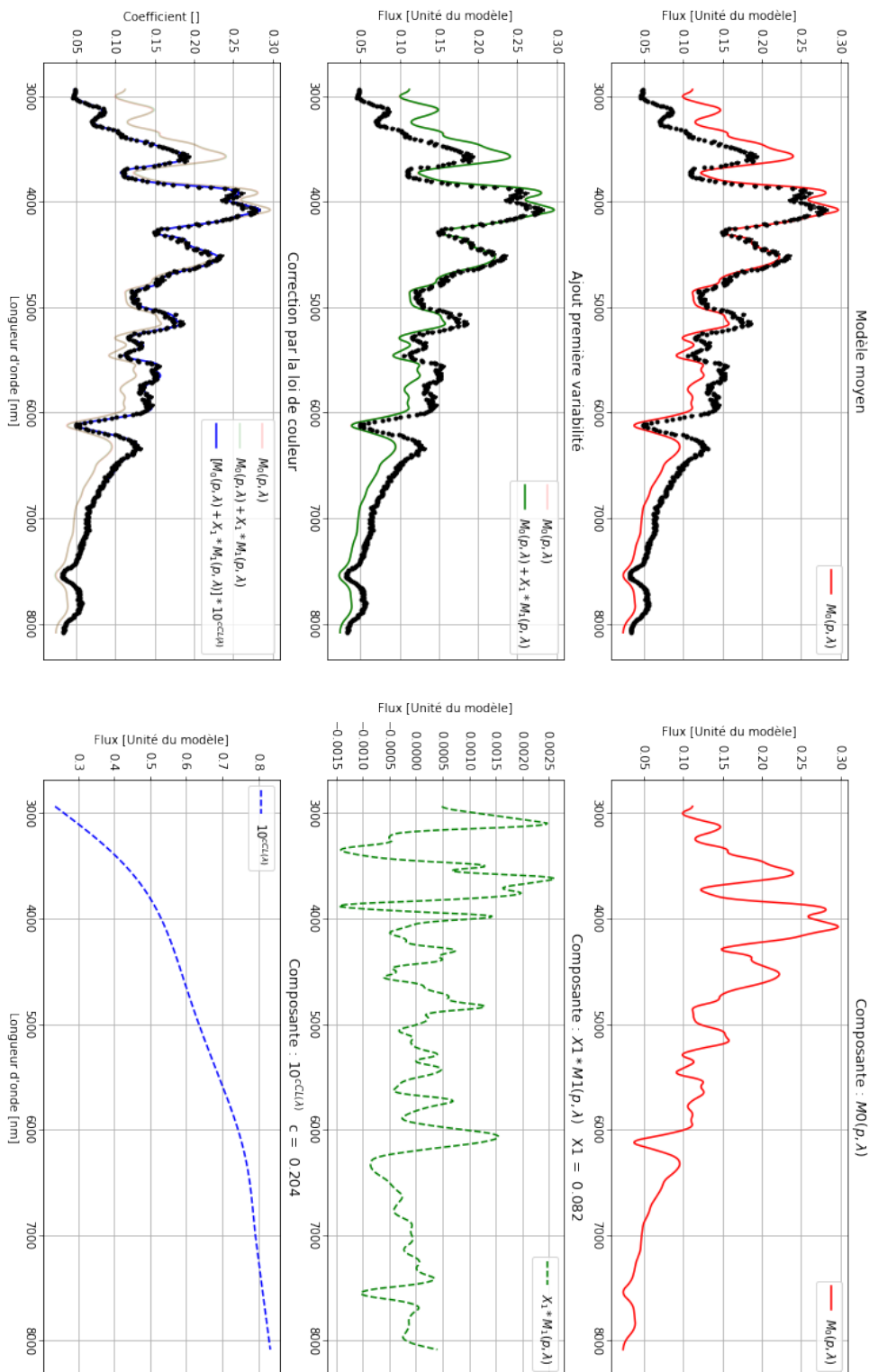


FIGURE 5.4 – Visualisation des différentes composantes spectrales. La colonne de gauche l'effet de la prise en compte successive des contributions de M_0 , M_1 et CL pour modéliser un spectre (points noir). Sur la colonne de droite, on peut voir la valeur des différentes contributions indépendamment.

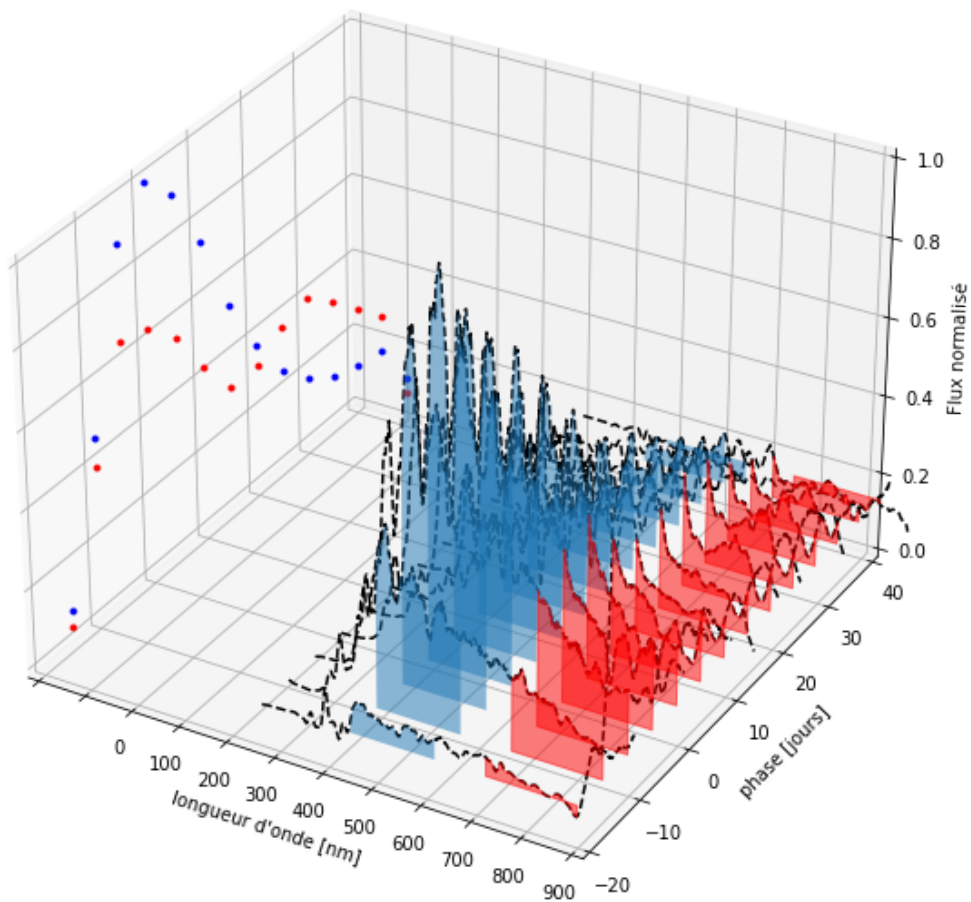


FIGURE 5.5 – Représentation de la modélisation de courbes de lumières. Création de deux courbes de lumière générées à travers deux filtres g (bleu) et i (rouge) de l'instrument du futur télescope du l'observatoire Vera Rubin.

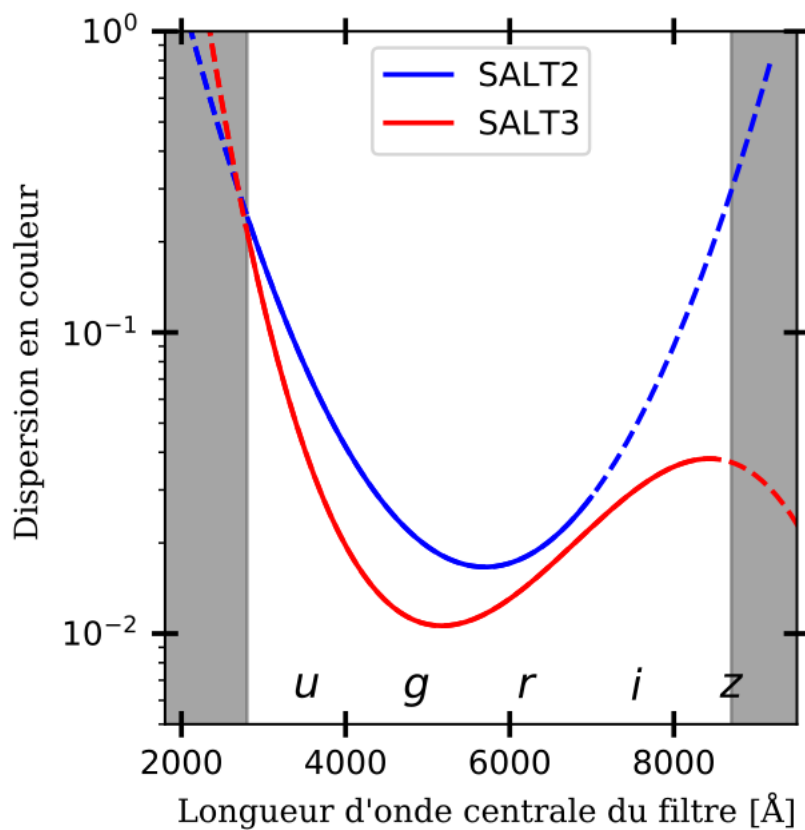


FIGURE 5.6 – Dispersion en couleur modélisé par SALT2 (en bleu) et SALT3 (en rouge). La paramétrisation du terme de dispersion en couleur, dépend du lot d'entraînement. Le modèle SALT2 paramétrise cette variance additionnel par un polynôme de degré 2, le modèle SALT3 quand à lui, par un polynôme de degré 3. Figure extraite de Kenworthy et al. (2021)

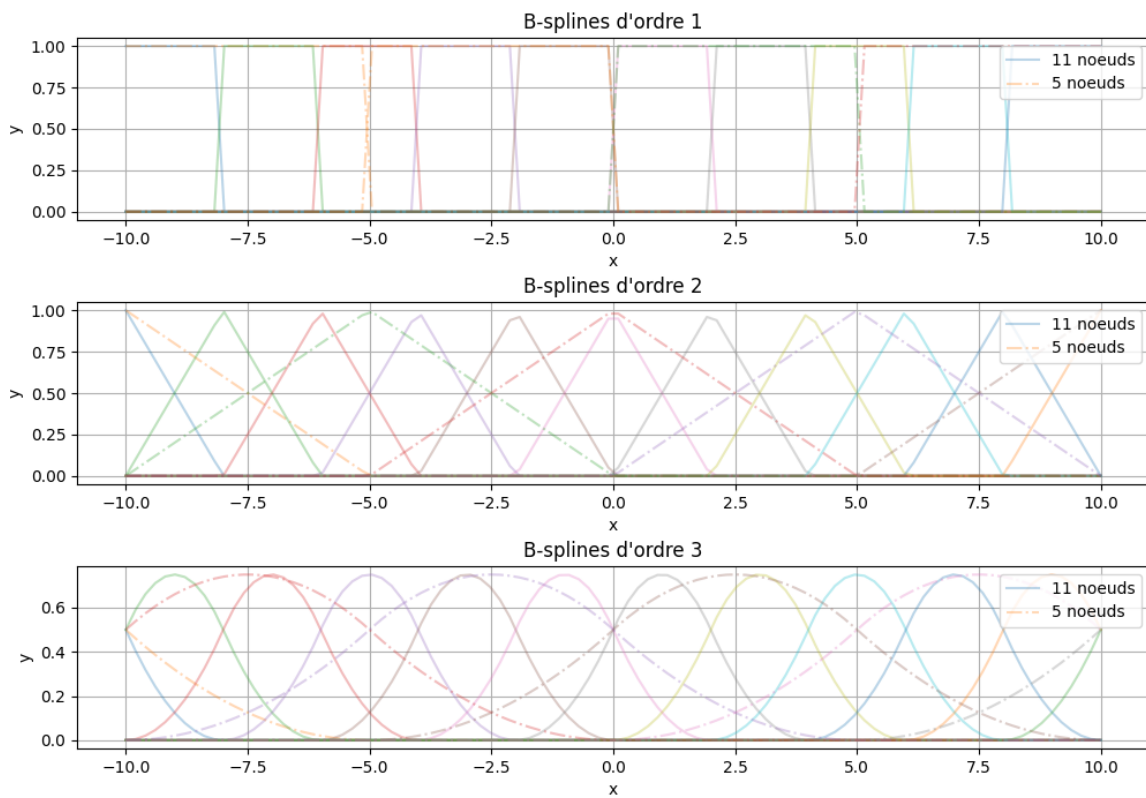


FIGURE 5.7 – Bases de B-splines d'ordre 1 (graphique du haut), d'ordre 2 (graphique du milieu) et d'ordre 3 (graphique du bas), avec 11 (courbes pleines) et 5 (courbes en tirets) nœuds. Plus l'ordre augmente, plus il y a de recouvrement entre les splines. Ici les graphiques ont été réalisés avec des paramètres tous égaux à un. Plus le nombre de nœud est grand, plus complexes sont les formes pouvant être décrites.

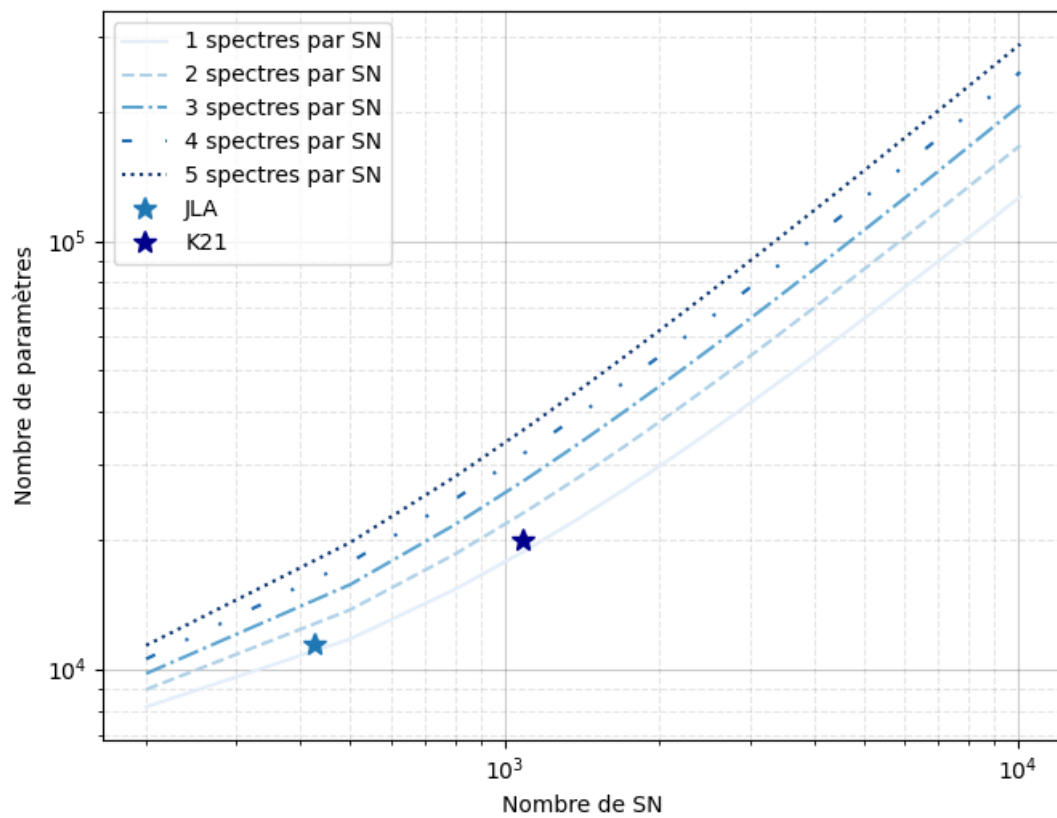


FIGURE 5.8 – Nombre de paramètres du modèle en fonction du nombre de SN. Le nombre de spectres par SN varie entre 1 et 5. Les deux étoiles correspondent aux nombres de SNe et de spectres contenus dans les lots d’entraînement de SALT2.4 (JLA) et SALT3 (K21).

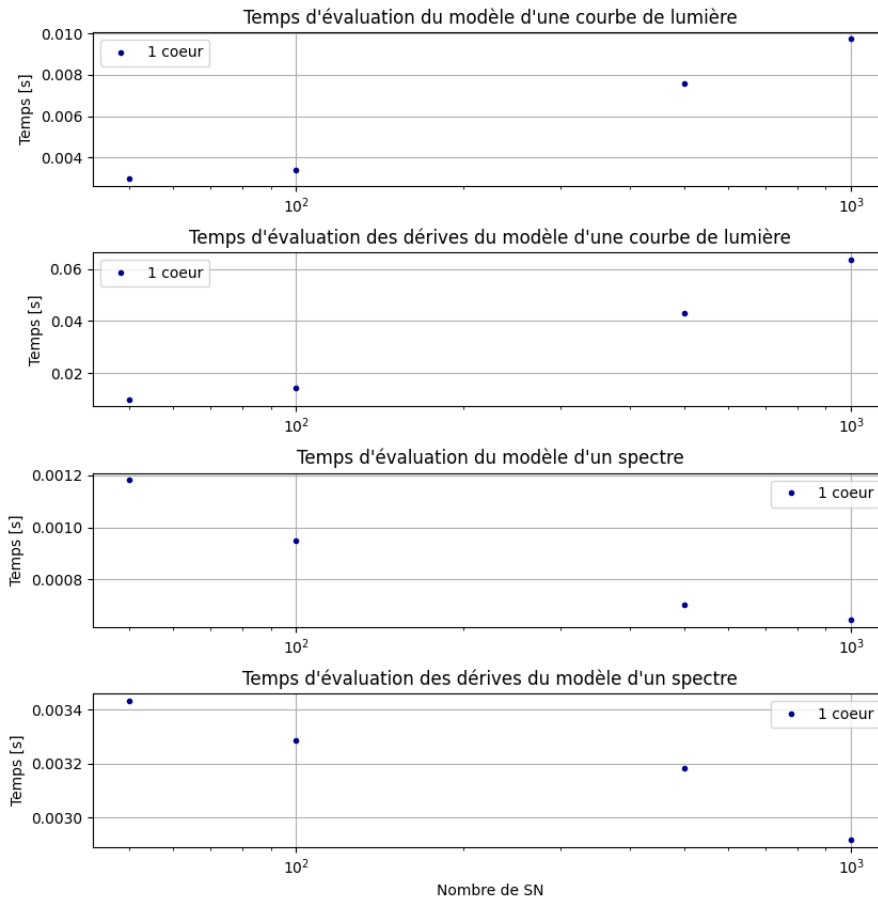


FIGURE 5.9 – Temps d'évaluation moyen, du modèle d'une courbe de lumière (panneau du haut), des dérivées du modèle d'une courbe de lumière (2ème panneau), du modèle d'un spectre (3ème panneau) et des dérivées du modèle d'un spectre (dernier panneau) en fonction du nombre de SNe dans le lot d'entraînement.

Minimiseur

Dans ce chapitre, nous détaillons la méthode développée pour réaliser l’ajustement non-linéaire sous contraintes, assorti d’un peu de régularisation, du modèle et du modèle d’erreur décrits dans le chapitre 5. Classiquement, l’approche consiste à construire, à partir du modèle et des données une fonction de vraisemblance $L(\theta)$ dont nous cherchons à déterminer le maximum. En pratique, nous minimisons la fonction $-2 \ln L$ à l’aide d’un algorithme du type Gauss-Newton. Dans une première partie, nous effectuons quelques rappels sur la méthode de maximum de vraisemblance. L’algorithme de minimisation est ensuite explicité. Dans la section suivante nous décrivons les dérivées premières et secondes de notre fonction de vraisemblance, et nous construisons l’équation normale dont la résolution permet de déterminer chaque pas de minimisation. Enfin nous décrivons l’algèbre de la propagation des erreurs et nous concluons.

Contents

6.1	Estimateur par maximum de vraisemblance.	144
6.2	Algorithme de score de Fisher : ajustement d’un modèle non linéaire	145
6.2.1	... sous contraintes	147
6.2.2	... avec de la régularisation	147
6.2.3	... en propageant les incertitudes de calibration	150
6.2.4	... et estimation de modèles d’erreur simultanément	151
6.2.5	χ^2 total	152
6.3	Les équations normales	153
6.3.1	Dérivées de l’ajustement du modèle	153
6.3.2	Dérivé du χ_{cons}^2	154
6.3.3	Dérivé du χ_{Reg}^2	155
6.3.4	Dérivé du χ_{calib}^2	155
6.3.5	Dérivé du χ_{CS}^2	155
6.3.6	L’équation finale	156
6.4	Estimation des incertitudes sur les paramètres	157
6.5	Conclusion	158

6.1 Estimateur par maximum de vraisemblance.

Nous groupons les N mesures dont nous disposons (photométriques et spectroscopiques) en un unique vecteur, noté \mathbf{Y} . Nous notons M le modèle, et nous groupons ses paramètres en un vecteur $\boldsymbol{\beta}$, de taille n . Les prédictions du modèle pour une valeur donnée de $\boldsymbol{\beta}$ sont notées $\mathbf{M}(\boldsymbol{\beta})$. La différence entre les mesures et les prédictions du modèle est notée $\mathbf{R} = \mathbf{Y} - \mathbf{M}(\boldsymbol{\beta})$ et est appelée vecteur des résidus. Les mesures sont affectées d'un bruit \mathbf{n} , supposé gaussien, dont nous notons \mathbf{V} la matrice de covariance :

$$\mathbf{n} \sim \mathcal{N}(0, \mathbf{V}) \quad (6.1)$$

Si le modèle décrit parfaitement les données, alors il existe une valeur de $\boldsymbol{\beta}$ notée $\boldsymbol{\beta}_{\text{true}}$, telle que :

$$\mathbf{Y} = \mathbf{M}(\boldsymbol{\beta}_{\text{true}}) + \mathbf{n} \quad (6.2)$$

Nous n'avons pas directement accès à $\boldsymbol{\beta}_{\text{true}}$. Nous pouvons l'estimer à partir d'une réalisation \mathbf{Y} des mesures. De nombreux estimateurs peuvent être imaginés. Celui ayant le plus de propriétés intéressantes est l'estimateur du maximum de vraisemblance. Il est construit de la manière suivante : la probabilité jointe d'observer \mathbf{Y} sachant $\boldsymbol{\beta}$ est appelée fonction de vraisemblance, et notée L . La valeur de $\boldsymbol{\beta}$ qui maximise $L(\mathbf{Y}|\boldsymbol{\beta})$ est appelée estimateur du maximum de vraisemblance de $\boldsymbol{\beta}_{\text{true}}$ et notée $\hat{\boldsymbol{\beta}}$.

L'estimateur du maximum de vraisemblance est consistant ($\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_{\text{true}} \rightarrow 0$ quand $N \rightarrow \infty$). Si le modèle est linéaire, $\hat{\boldsymbol{\beta}}$ est un estimateur de $\boldsymbol{\beta}_{\text{true}}$ qui est non biaisé et efficace, au sens qu'il atteint la borne de variance minimale. Ces propriétés tombent si le modèle est non linéaire. En revanche, on peut montrer que dans ce cas, $\hat{\boldsymbol{\beta}}$ est asymptotiquement non biaisé et asymptotiquement efficace (i.e. non biaisé et efficace dans la limite où $N \rightarrow \infty$). Enfin, si l'on peut construire une expression analytique pour la vraisemblance, l'estimateur est relativement simple à mettre en œuvre, pour peu qu'on ait sous la main un code de maximisation efficace.

Dans notre cas particulier, le bruit affectant les mesures étant gaussien, L s'écrit simplement :

$$L(\mathbf{Y}|\boldsymbol{\beta}) = \frac{1}{(2\pi)^{N/2} |\mathbf{V}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{Y} - \mathbf{M}(\boldsymbol{\beta}))^T \mathbf{W} (\mathbf{Y} - \mathbf{M}(\boldsymbol{\beta})) \right] \quad (6.3)$$

où $\mathbf{W} = \mathbf{V}^{-1}$ est l'inverse de la matrice de covariance, encore appelée matrice des poids dans la suite.

En pratique, on préfère généralement minimiser la quantité $-2 \ln L$. On obtient donc, en posant $\mathbf{R}(\boldsymbol{\beta}) = \mathbf{Y} - \mathbf{M}(\boldsymbol{\beta})$:

$$-2 \ln(L(\boldsymbol{\beta})) = N \ln(2\pi) + \ln(|\mathbf{V}|) + \mathbf{R}^T \mathbf{V}^{-1} \mathbf{R} \quad (6.4)$$

Si la matrice de covariance \mathbf{V} ne dépend pas des paramètres, par exemple, si elle ne contient que des erreurs de mesure connues, alors, minimiser $-2 \ln L$ revient à minimiser la fonction χ^2 classique :

$$\chi^2 = \mathbf{R}^T \mathbf{V}^{-1} \mathbf{R} \quad (6.5)$$

On retrouve les moindres carrés ordinaires. En revanche, si la matrice de covariance dépend d'une partie des paramètres, par exemple des paramètres du modèle, ou de paramètres spécifiques au modèle d'erreur, alors il est nécessaire d'inclure le terme $\ln |\mathbf{V}|$ dans la minimisation.

6.2 Algorithme de score de Fisher : ajustement d'un modèle non linéaire

Dans un premier temps, nous nous intéressons à la minimisation d'un χ^2 ordinaire, tel que défini équation 6.5. Minimiser le χ^2 consiste à trouver la valeur $\hat{\beta}$ qui annule son gradient.

Modèles linéaires Si le modèle est linéaire, il peut s'écrire sous la forme :

$$\mathbf{M}(\beta) = \mathbf{J}\beta \quad (6.6)$$

où \mathbf{J} est la matrice $N \times n$ des dérivées du modèle par rapport aux paramètres. Le χ^2 est alors une fonction quadratique de β . Son gradient est une fonction linéaire qui peut s'écrire :

$$\nabla(\chi^2) = 2\mathbf{J}^T\mathbf{W}\mathbf{J}\beta - 2\mathbf{J}^T\mathbf{W}\mathbf{Y} \quad (6.7)$$

Et le minimum du χ^2 s'obtient facilement en résolvant le système linéaire ci-dessus (aussi appelé équation normale du problème) :

$$\hat{\beta} = \beta_{\min} = (\mathbf{J}^T\mathbf{W}\mathbf{J})^{-1} \mathbf{J}^T\mathbf{W}\mathbf{Y} \quad (6.8)$$

Modèles non-linéaires Si le modèle est non-linéaire, alors la détermination de son minimum se fait de manière itérative, à partir d'une estimation initiale β_0 de la solution en linéarisant le modèle à chaque étape :

$$\mathbf{M}(\beta_i + \delta\beta) = \mathbf{M}(\beta_i) + \mathbf{J}\delta\beta \quad \text{avec} \quad \mathbf{J} = \begin{pmatrix} \frac{\partial m_1}{\partial \beta_1} & \cdots & \frac{\partial m_1}{\partial \beta_n} \\ \vdots & & \vdots \\ \frac{\partial m_N}{\partial \beta_1} & \cdots & \frac{\partial m_N}{\partial \beta_n} \end{pmatrix} \quad (6.9)$$

À l'étape i :

$$\chi^2(\beta_i + \delta\beta) = \chi^2(\beta_i) + \nabla\chi^2(\beta_i) \cdot \delta\beta + \frac{1}{2}\delta\beta^T\mathbf{H}\delta\beta \quad (6.10)$$

où \mathbf{H} est la matrice des dérivées secondes du χ^2 , appelée matrice hessienne. Le pas est la solution de l'équation normale ci-dessus :

$$\delta\beta = -\mathbf{H}^{-1}(\beta_i) \nabla\chi^2(\beta_i) \quad (6.11)$$

et on met à jour $\beta_{i+1} = \beta_i + \delta\beta$. Cette technique est appelée algorithme de Newton. C'est l'un des algorithmes les plus rapides pour trouver le minimum d'une fonction. Il est cependant assez sensible à la qualité des conditions initiales.

Nous avons donc besoin de calculer à chaque itération, le gradient et la hessienne de l'équation 6.5, dérivées premières et secondes par rapport aux paramètres β :

La dérivée du χ^2 par rapport au paramètre β_i s'écrit :

$$\frac{\partial\chi^2}{\partial\beta_i} = -2\frac{\partial\mathbf{M}(\beta)^T}{\partial\beta_i} \mathbf{V}^{-1}\mathbf{R} \quad (6.12)$$

Pour le gradient $\nabla\chi^2$:

$$\nabla\chi^2 = -2\mathbf{J}^T\mathbf{V}^{-1}\mathbf{R} \quad (6.13)$$

Quant à la hessienne, nous l'obtenons en dérivant l'équation 6.12 :

$$\frac{\partial^2\chi^2}{\partial\beta_i\partial\beta_j} = -2\frac{\partial^2\mathbf{M}(\beta)^T}{\partial\beta_i\partial\beta_j}\mathbf{V}^{-1}\mathbf{R} + 2\frac{\partial\mathbf{M}(\beta)^T}{\partial\beta_i}\mathbf{V}^{-1}\frac{\partial^2\mathbf{M}(\beta)}{\partial\beta_j} \quad (6.14)$$

En notant $\mathbf{H}_{\chi^2} = \left(\frac{\partial^2\chi^2}{\partial\beta_i\partial\beta_j}\right)$, le hessien du χ^2 et $\mathbf{H}_M = \left(\frac{\partial^2\mathbf{M}(\beta)}{\partial\beta_i\partial\beta_j}\right)$ le hessien du modèle, on peut réécrire l'équation 6.14.

$$\mathbf{H}_{\chi^2} = -2\mathbf{H}_M^T\mathbf{V}^{-1}\mathbf{R} + 2\mathbf{J}^T\mathbf{V}^{-1}\mathbf{J} \quad (6.15)$$

En pratique, l'algorithme de Newton présente quelques limitations. D'abord, le calcul des dérivées secondes du modèle peut être laborieux, chronophage et gourmand en mémoire. Elles ont de plus une contribution négligeable et peuvent causer des instabilités de convergence. Nous utilisons en pratique un cas particulier, appelé algorithme de Gauss-Newton ou score de Fisher¹ [Fisher (1992)], pour lequel il convient de remplacer le hessien dans l'équation 6.17 par son espérance :

$$\mathbb{E}[\mathbf{H}_{\chi^2}] = \mathbb{E}[-2\mathbf{H}_M^T\mathbf{V}^{-1}\mathbf{R} + 2\mathbf{J}^T\mathbf{V}^{-1}\mathbf{J}] = 2\mathbf{J}^T\mathbf{V}^{-1}\mathbf{J} \quad (6.16)$$

On obtient donc finalement pour les équation normales :

$$\delta\beta = -(\mathbb{E}[\mathbf{H}_{\chi^2}])^{-1}\nabla\chi^2 \quad (6.17)$$

soit :

$$\delta\beta = (\mathbf{J}^T\mathbf{V}^{-1}\mathbf{J})^{-1}\mathbf{J}^T\mathbf{V}^{-1}\mathbf{R} \quad (6.18)$$

Le minimiseur que nous avons développé pour notre modèle s'articule autour de l'algorithme que nous venons de décrire. Il s'agit, à notre connaissance, de la manière la plus efficace pour trouver la solution d'un problème non-linéaire comprenant un grand nombre de paramètres. Chaque étape de l'itération requiert une évaluation du modèle et de ses dérivées, ainsi que le calcul et l'inversion de la hessienne : $\mathbf{H} = 2\mathbf{J}^T\mathbf{W}\mathbf{J}$. Comme la Hessienne est définie-positive, il est généralement avantageux de ne pas l'inverser explicitement, mais de se contenter d'une factorisation triangulaire, de type $\mathbf{L}\mathbf{L}^T$ (décomposition de Cholesky). Il est également fréquent que la Hessienne soit creuse. Dans ce cas, l'utilisation d'une version de l'algorithme de Cholesky adaptée aux matrices creuses (e.g. `cholmod`) peut accélérer significativement chaque itération.

1. Cas particulier d'un algorithme de Gauss Newton, appliqué à un ajustement par maximum de log-vraisemblance.

6.2.1 ... sous contraintes

Pour briser les dégénérescences entre les paramètres du modèle, nous imposons sept contraintes (cf. section 5.2.3). Les six premières sont linéaires, impliquant un formalisme simple. On peut les écrire sous la forme du produit matriciel suivant : $\mathbf{H}_{\text{pen}}\boldsymbol{\beta} = \boldsymbol{\alpha}$, où \mathbf{H}_{pen} est une matrice dont chaque ligne comporte les coordonnées de paramètres à contraindre. La valeur de la contrainte est stockée dans le vecteur $\boldsymbol{\alpha}$. Par exemple, si nous avons 200 SNe, avec chacune 1 spectres et 4 courbes de lumière, nous avons au total 10484 paramètres, \mathbf{H}_{pen} sera une matrice 10484×6 , la sixième ligne correspondant à la contraintes sur la moyenne de couleur c , elle sera seulement de un aux indices correspondant aux paramètres c dans $\boldsymbol{\beta}$. $\boldsymbol{\alpha}$ sera un vecteur de 6 valeurs. Suivant notre exemple, son dernier terme sera nulle puisque la moyenne des couleurs est nulle. On ajoute à l'équation 6.5, les contraintes sous forme d'une pénalité quadratique, minimisant la distance entre $\mathbf{H}_{\text{pen}}\boldsymbol{\beta}$ et $\boldsymbol{\alpha}$.

En revanche, la contrainte sur la variance des \mathbf{X}_1 , est non linéaire. Comme la valeur moyenne des \mathbf{X}_1 est nulle (c'est une des contraintes linéaires du problème), nous pouvons l'écrire sous la forme :

$$\frac{1}{N} \mathbf{X}_1^T \mathbf{X}_1 = 1 \quad (6.19)$$

Le paramètre d'intensité de cette contrainte n'est pas le même que le paramètre pour les contraintes linéaires.

L'ensemble des contraintes est noté sous la forme : $\mathbf{C}(\boldsymbol{\beta})$, qui est égal à la distance entre la contrainte sur les paramètres et la valeur attendue. Le vecteur des intensités des contraintes linéaires, noté μ_{pen} , et celui des contraintes non-linéaires, noté μ_{pen}^1 sont regroupés dans le vecteur μ_{pen} . On ajoute donc au χ^2 le terme suivant :

$$\chi_{\text{cons}}^2 = \mu_{\text{pen}} \mathbf{C}(\boldsymbol{\beta})^T \mathbf{C}(\boldsymbol{\beta}) \quad (6.20)$$

Ce dernier terme dans l'équation 6.20 peut être vu comme un a priori gaussien sur la distribution des paramètres impliqués dans les contraintes. Il impose aux contraintes \mathbf{C} d'être nulles avec une précision de l'ordre de $\sqrt{1/\mu_{\text{pen}}}$. On note que toutes ces contraintes brisent des dégénérescences exactes entre les paramètres du modèle. On vérifie en pratique que la contribution de χ_{cons}^2 au χ^2 total est négligeable, et que les contraintes sont parfaitement respectées à partir du moment où la valeur de μ_{pen} est assez élevée (quelques 10^{10}).

6.2.2 ... avec de la régularisation

La régularisation a pour but de contraindre les surfaces M_0 et M_1 dans les zones du plan (λ, p) vides de données (cf. section 5.2.5). Nous avons recours à un schéma de régularisation classique, appelé *Régularisation de Tikhonov généralisée*, consistant à ajouter au χ^2 (6.5) une petite pénalité quadratique. Cette pénalité est conçue pour limiter les variations brusques du modèle (oscillations) et faire en sorte que le modèle tende lentement vers zéro en l'absence de données, en particulier aux phases $p < -20$ jours et $p > 60$ jours. Pour le moment, la régularisation porte sur les variations du modèle en phase. La densité de spectres permet de contraindre correctement la base, et la densité de points sur les spectres semble limiter les phénomènes d'oscillation. La pénalité de régularisation est du type :

$$\chi_{reg}^2 = \mu_{reg} \boldsymbol{\beta}^T \mathbf{P} \boldsymbol{\beta} \quad (6.21)$$

avec \mathbf{P} , la matrice de régularisation dont la forme permet de contraindre les paramètres $\theta_{0|1}$. La régularisation souhaitée porte sur les valeurs des paramètres θ et leurs variations de proche en proche :

$$\chi_{reg}^2 = \mu_{reg} \left[\sum_{i=0}^{N_\lambda-1} \sum_{i=0}^{N_{phase}-1} \left(\theta_{0|1}^{j,i+1} - \theta_{0|1}^{j,i} \right)^2 + \left(\theta_{0|1}^{j,i} \right)^2 \right] \quad (6.22)$$

Le vecteur $\boldsymbol{\beta}$ est organisé de manière à ce que les paramètres des surfaces, $\theta_{0|1}$ sont alignées par phase pour chaque longueur d'onde. Ainsi la matrice, \mathbf{P} prend la forme suivante, non nulle pour les paramètres $\theta_{0|1}$, en listant les paramètres en quatre blocs ($[X_0, X_1, c, t_{max}^B], M_0, M_1, [CL, s, \eta, \kappa]$) :

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \mathbf{P}_{\theta_{0|1}} & 0 & 0 \\ 0 & 0 & \mathbf{P}_{\theta_{0|1}} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (6.23)$$

avec :

$$\mathbf{P}_{\theta_{0|1}} = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 3 & -1 & \dots & 0 \\ \dots & & & & \dots \\ \dots & & -1 & 3 & -1 & \dots \\ \dots & & & & & \dots \\ 0 & & & & -1 & 3 & -1 \\ 0 & & & & & -1 & 2 \end{pmatrix} \quad (6.24)$$

μ_{pen} est un hyper-paramètre dont la valeur doit être ajustée de manière à ne pas biaiser l'estimation des paramètres $\boldsymbol{\beta}$. À ce stade de développement du modèle, nous utiliserons des simulations Monte-Carlo pour en déterminer une valeur raisonnable.

À des fins d'illustration, nous présentons figure 6.1 l'impact du paramètre μ_{reg} sur l'ajustement du modèle. Ce sont les résultats de l'ajustement de splines, définis sur l'intervalle $[-100, 100]$ jours, sur deux courbes de lumière comportant 8 points (colonne de gauche) et 50 points (colonne de droite). Chaque ligne correspond à différentes valeurs de régularisation. Les courbes rouges correspondent à une matrice \mathbf{P} égale à l'identité contraignant seulement les paramètres non-contraints et non les variations en paramètres de phase consécutifs. Les courbes bleues correspondent à un ajustement avec une matrice contraignant les dérivées entre les paramètres. On peut voir sur le premier graphe en haut à gauche que lorsque une courbe de lumière n'est pas bien échantillonnée, il peut y avoir des oscillations de splines entre les points, comme on peut le voir à environ +25 jours. Lorsqu'une courbe de lumière est mieux échantillonnée, ces oscillations disparaissent. Au passage des données à l'absence de données, les splines peuvent aussi présenter des oscillations, dues à une contrainte partielle sur les coefficients. Une augmentation de l'intensité des contraintes permet d'éviter ce phénomène. La ligne du milieu montre que si

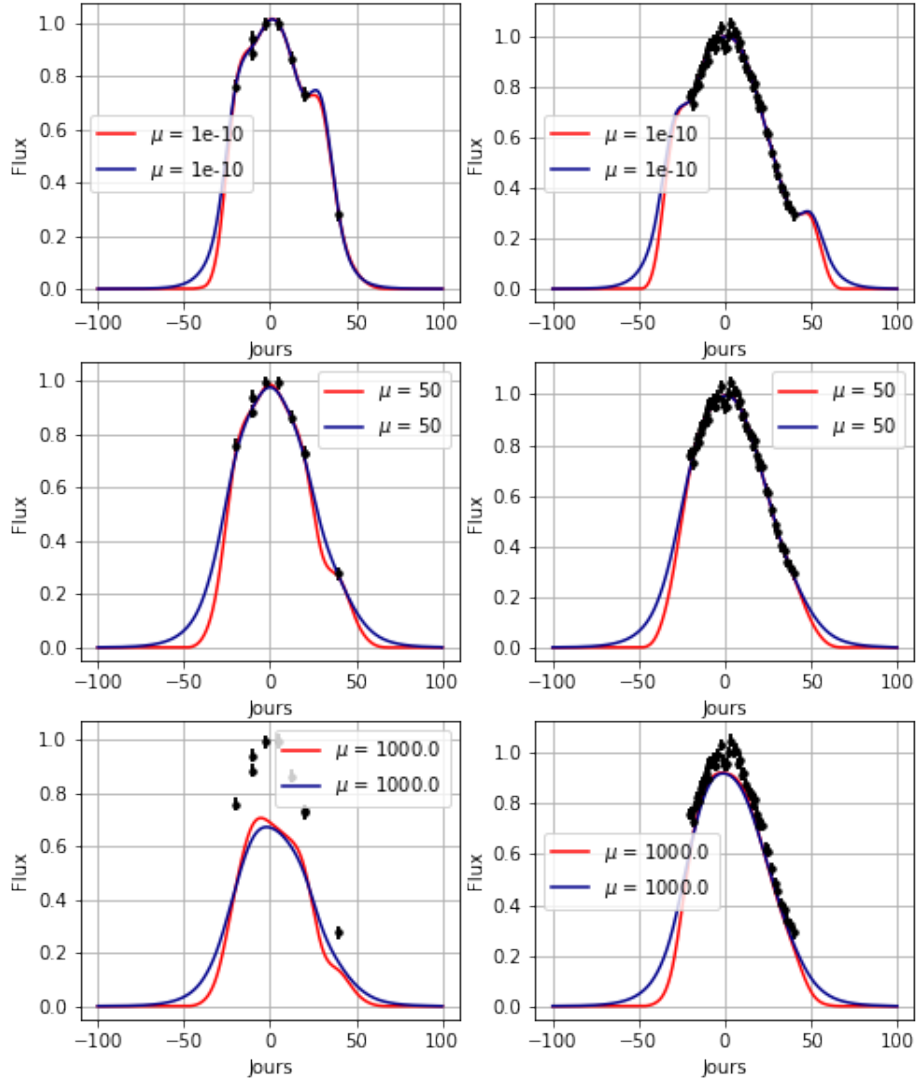


FIGURE 6.1 – Impact du coefficient de régularisation, μ_{reg} , sur l'ajustement d'une courbe de lumière par des splines. Les points noirs correspondent à des courbes de lumière simulées avec différents échantillonnages, 8 points à gauche et 50 points à droites. Pour les modèles rouges la régularisation pénalise les paramètres non contraints. Pour les modèles bleus, la régularisation pénalise aussi la pente entre les paramètres consécutifs en phase.

l'hyper-paramètre μ est bien choisi alors le modèle est lissé et les artefacts vus au-dessus disparaissent. On peut voir avec le modèle rouge de petites oscillations résiduelles, complètement retirées par le modèle bleu.

Naïvement, prendre un coefficient très élevé impliquerait une contraintes toujours très forte

et constamment respectée. Or nous pouvons voir sur la troisième ligne de la figure 6.1 que si la contrainte est trop importante, alors on introduit un biais dans l'ajustement et le modèle ne décrit plus les données. Le point important est donc de choisir μ_{reg} en fonction du lot d'entraînement, assez grand pour qu'il puisse régulariser le modèle mais pas trop pour ne pas introduire de biais. Le choix de ce paramètre sera discuté au chapitre 8.

6.2.3 ... en propageant les incertitudes de calibration

Au premier ordre, les incertitudes de calibration génèrent des modes d'erreur spécifiques à chaque bande du relevé, chaque mode affectant *toutes* les mesures prises dans la bande correspondante. On peut considérer l'impact de la calibration comme un bruit additionnel et modifier l'équation 6.2 comme suit :

$$\mathbf{Y} = \mathbf{M}(\boldsymbol{\beta}_{\text{true}}) + \mathbf{n}_{\text{cal}} + \mathbf{n}_{\text{meas}} \quad (6.25)$$

\mathbf{n}_{cal} est le bruit de calibration, et \mathbf{n}_{meas} est le bruit de mesure (noté \mathbf{n} dans l'équation 6.2). Le bruit de mesure est supposé non corrélé ; la matrice de covariance associée \mathbf{V}_{meas} est diagonale et de rang N . En revanche, si nous avons n_b bandes, alors nous aurons au plus le même nombre de modes d'erreur de calibration indépendants. Le rang de la matrice de covariance \mathbf{V}_{cal} du vecteur \mathbf{n}_{cal} est donc au plus n_b . Pour faire apparaître ceci explicitement, on peut écrire :

$$\mathbf{n}_{\text{cal}} = \mathbf{K} \boldsymbol{\eta} \quad \text{avec} \quad \boldsymbol{\eta} \sim \mathcal{N}(0, \mathbf{V}_{\boldsymbol{\eta}}) \quad (6.26)$$

où $\boldsymbol{\eta}$ est un vecteur aléatoire de dimension n_b , \mathbf{K} est une matrice de mélange $N \times n_b$ propageant l'impact du bruit de calibration dans chaque bande sur les mesures, et $\mathbf{V}_{\boldsymbol{\eta}}$ est la matrice de covariance des $\boldsymbol{\eta}$, autrement dit, la matrice de covariance des paramètres de calibration.

Pour prendre en compte les incertitudes de calibration, il nous faut donc modifier la vraisemblance présentée équation 6.3 en substituant à \mathbf{V}_{meas} la matrice de covariance totale $\mathbf{V}_{\text{meas}} + \mathbf{V}_{\text{cal}} = \mathbf{V}_{\text{meas}} + \mathbf{K}\mathbf{V}_{\boldsymbol{\eta}}\mathbf{K}^T$.

La fonction que nous avons à minimiser devient :

$$-2 \ln L = \mathbf{R}^T (\mathbf{V}_{\text{meas}} + \mathbf{V}_{\text{cal}})^{-1} \mathbf{R} + \dots \quad (6.27)$$

Minimiser directement cette log-vraisemblance est nettement plus difficile. En effet, la matrice de covariance est non seulement de grandes dimensions ($N \times N$ avec $N \sim O(10^4 - 10^5)$) mais également dense. Ce qui était faisable avec une matrice diagonale, ne l'est plus avec une matrice dense, sauf à disposer de machines à très grande mémoire, et de beaucoup de temps de calcul.

On peut remarquer cependant que la vraisemblance jointe $L(\mathbf{Y}, \boldsymbol{\eta} | \boldsymbol{\beta})$ que nous cherchons à maximiser peut s'exprimer simplement comme :

$$L(\mathbf{Y}, \boldsymbol{\eta} | \boldsymbol{\beta}) = L(\mathbf{Y} | \boldsymbol{\beta}, \boldsymbol{\eta}) \times p(\boldsymbol{\eta}) \quad (6.28)$$

Autrement dit, nous pouvons aussi considérer les $\boldsymbol{\eta}$ comme n_b paramètres additionnels du modèle et maximiser la fonction plus simple $L(\mathbf{Y} | \boldsymbol{\beta}, \boldsymbol{\eta})$ assortie d'un prior sur les $\boldsymbol{\eta}$. Cet a priori est connu, puisque : $\boldsymbol{\eta} \sim \mathcal{N}(0, \mathbf{V}_{\boldsymbol{\eta}})$.

Notre problème consiste donc à minimiser une quantité de la forme :

$$-2 \ln L = (\mathbf{Y} - \mathbf{J}\boldsymbol{\beta} - \mathbf{K}\boldsymbol{\eta})^T \mathbf{V}_{\text{meas}}^{-1} (\mathbf{Y} - \mathbf{J}\boldsymbol{\beta} - \mathbf{K}\boldsymbol{\eta}) + \boldsymbol{\eta}^T \mathbf{V}_{\boldsymbol{\eta}}^{-1} \boldsymbol{\eta} \quad (6.29)$$

par rapport aux paramètres $(\boldsymbol{\beta}, \boldsymbol{\eta})$, (puis marginaliser sur $\boldsymbol{\eta}$), plutôt que minimiser directement la log-vraisemblance initiale par rapport aux seuls paramètres $\boldsymbol{\beta}$:

$$-2 \ln L = (\mathbf{Y} - \mathbf{J}\boldsymbol{\beta})^T (\mathbf{V}_{\text{meas}} + \mathbf{K}\mathbf{V}_{\boldsymbol{\eta}}\mathbf{K}^T)^{-1} (\mathbf{Y} - \mathbf{J}\boldsymbol{\beta}) \quad (6.30)$$

L'équation normale associée à notre nouvelle vraisemblance 6.29 est :

$$\begin{pmatrix} \mathbf{J}^T \mathbf{V}_{\text{meas}}^{-1} \mathbf{J} & \mathbf{J}^T \mathbf{V}_{\text{meas}}^{-1} \mathbf{K} \\ \mathbf{K}^T \mathbf{V}_{\text{meas}}^{-1} \mathbf{J} & \mathbf{K}^T \mathbf{V}_{\text{meas}}^{-1} \mathbf{K} + \mathbf{V}_{\eta}^{-1} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\eta} \end{pmatrix} = \begin{pmatrix} \mathbf{J}^T \mathbf{V}_{\text{meas}}^{-1} \mathbf{Y} \\ \mathbf{K}^T \mathbf{V}_{\text{meas}}^{-1} \mathbf{Y} \end{pmatrix} \quad (6.31)$$

On peut montrer, que l'équation normale issue de la vraisemblance initiale 6.30 est très exactement la même que celle obtenue de cette nouvelle vraisemblance 6.29 après avoir substitué $\boldsymbol{\eta}$ (cela requiert juste d'utiliser la formule de Woodbury $(\mathbf{A} + \mathbf{C}\mathbf{B}\mathbf{C}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{B}^{-1} + \mathbf{C}^T\mathbf{B}^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{A}^{-1}$ pour faire l'identification).

Nous avons donc un formalisme pour propager les incertitudes de calibration, et en pratique, tout jeu d'incertitudes corrélées dont le rang est faible devant N , en ajoutant au modèle un jeu de paramètres, tenus par un prior. Ce type de formalisme est appelé modèle mixte. Il est inspiré des modèles linéaires mixtes proposés par [Henderson et al. (1959)].

In fine, on décide donc de transformer le modèle pour qu'il incorpore l'effet de la calibration, et on ajoute au χ^2 un terme :

$$\chi_{\text{calib}}^2(\boldsymbol{\beta}) = \boldsymbol{\eta}^T \mathbf{V}_{\eta}^{-1} \boldsymbol{\eta} \quad (6.32)$$

6.2.4 ... et estimation de modèles d'erreur simultanément

L'estimation de la variance simultanément avec celle du modèle est plus délicate algébriquement. Elle conduit généralement à un estimateur biaisé, dont le biais doit être évalué sur des simulations.

Error snake Nous modélisons la variance additionnelle à l'aide d'une matrice diagonale \mathbf{V}_{mod} dont les termes diagonaux sont définis dans l'équation 5.29. La matrice de poids est donc :

$$\mathbf{V}^{-1} = (\mathbf{V}_{\text{meas}} + \mathbf{V}_{\text{mod}})^{-1} \quad (6.33)$$

\mathbf{V}_{mod} dépend d'un jeu de paramètres spécifiques ($\boldsymbol{\gamma}$) qui seront inclus dans l'ajustement, ainsi que d'un sous-ensemble des paramètres $\boldsymbol{\beta}$ du modèle. Par conséquent, le terme $\ln|\mathbf{V}|$ présent dans l'expression de la log-vraisemblance n'est plus constant et ses dérivées doivent être évaluées et incluses dans la minimisation. Le calcul de ces dérivées est assez coûteux, mais néanmoins faisable tant que la matrice des poids reste diagonale. Ce calcul est détaillé section 6.3.

On peut réécrire l'équation 6.5 :

$$\chi_{\text{model}}^2 = N \ln(2\pi) + \ln(|\mathbf{V}(\boldsymbol{\beta}, \boldsymbol{\gamma})|) + \mathbf{R}(\boldsymbol{\beta})^T \mathbf{V}^{-1}(\boldsymbol{\beta}, \boldsymbol{\gamma}) \mathbf{R}(\boldsymbol{\beta}) \quad (6.34)$$

Color scatter De même que les incertitudes de calibration, l'effet de la dispersion en couleur peut-être modélisé directement en ajoutant des termes non-diagonaux dans la matrice des poids. Afin de préserver une matrice de poids diagonale, nous choisissons d'utiliser un formalisme analogue à celui mis en oeuvre pour la propagation des incertitudes de calibration.

Pour modéliser le *color scatter*, i.e. la dispersion autour de la loi de couleur, nous introduisons dans l'ajustement un paramètre supplémentaire par bande par SN. Ces paramètres sont notés κ . Ils sont tenus par un a priori empirique, donc les paramètres sont également ajustés.

$$\kappa \sim \mathcal{N}(0, \mathbf{V}_\kappa(\sigma_\kappa)) \quad (6.35)$$

La matrice \mathbf{V}_κ est construite à partir d'un polynôme $P(\lambda_r)$ de la longueur d'onde *restframe*. À chaque courbe de lumière, on peut associer la longueur effective du filtre dans le référentiel de la SN λ_{eff} . La variance du terme κ associé est donnée par $P(\lambda_{\text{eff}})$. Les coefficients de P , de degré 3, sont notés $\sigma_{\kappa,i}$ et sont inclus dans l'ajustement.

On ajoute donc à la log-vraisemblance totale les termes suivants :

$$-2 \ln(f(\kappa)) = N_{lc} \ln(2\pi) + \ln(|\mathbf{V}_\kappa(\sigma_\kappa)|) + \kappa^T \mathbf{V}_\kappa(\sigma_\kappa)^{-1} \kappa \quad (6.36)$$

où N_{lc} est le nombre de courbes de lumières incluses dans le lot d'entraînement. La encore le dernier terme est une contrainte permettant de briser la dégénérescence entre les paramètres κ et les paramètres du modèles. Puisque nous avons un paramètre κ par courbe de lumière, ils seront dégénérés avec les X_0 , avec les surfaces, les paramètres de calibration. On ajoute ainsi deux nouveaux termes au χ^2 total :

$$\chi_{CS}^2(\kappa, \sigma_\kappa) = \ln |\mathbf{V}_\kappa(\sigma_\kappa)| + \kappa^T \mathbf{V}_\kappa^{-1}(\sigma_\kappa) \kappa \quad (6.37)$$

6.2.5 χ^2 total

Finalement, nous nous retrouvons donc dans le cas de devoir minimiser la fonction suivante :

$$\begin{aligned} \chi^2 &= \chi_{model}^2 + \chi_{cons}^2 + \chi_{reg}^2 + \chi_{calib}^2 + \chi_{CS}^2 \\ &= Cst + \ln(|\mathbf{V}(\beta, \gamma)|) + \mathbf{R}(\beta)^T \mathbf{V}^{-1}(\beta, \gamma) \mathbf{R}(\beta) \\ &\quad + \mu_{pen} \mathbf{C}(\beta)^T \mathbf{C}(\beta) \\ &\quad + \mu_{reg} \beta^T \mathbf{P} \beta \\ &\quad + \eta^T \mathbf{V}_\eta^{-1} \eta \\ &\quad + \ln |\mathbf{V}_\kappa(\sigma_\kappa)| + \kappa^T \mathbf{V}_\kappa^{-1}(\sigma_\kappa) \kappa \end{aligned} \quad (6.38)$$

Rien ne garantissait que cette minimisation puisse s'effectuer en une seule étape. Le développement du minimiseur s'est donc fait de proche en proche, en complexifiant peu à peu le modèle, et en s'appuyant sur des simulations de complexité croissante.

Dans la prochaine section, nous détaillons l'équation normale du problème complet.

6.3 Les équations normales

Dans cette section nous déterminons l'incrément à calculer à chaque itération, défini équation 6.17. Nous allons calculer les dérivées premières et secondes par rapport aux différents paramètres du modèle et du modèle d'erreur pour chaque terme de l'équation 6.38. Dans un souci de simplicité, nous allons regrouper tous les paramètres sous une même variable $\rho = \begin{pmatrix} \beta \\ \gamma \\ \sigma_\kappa \end{pmatrix}$ avec β étant les paramètres du modèle, γ les paramètres de l'*error snake* et σ_κ les paramètres du *color scatter*.

6.3.1 Dérivées de l'ajustement du modèle

On peut réécrire l'équation 6.34 en fonction des paramètres ρ et la dérivée par rapport au paramètre ρ_i :

$$\begin{aligned} \chi_{model}^2(\rho) &= N \ln(2\pi) + \ln(|\mathbf{V}(\rho)|) + \mathbf{R}(\rho)^T \mathbf{V}^{-1}(\rho) \mathbf{R}(\rho) \\ \frac{\partial \chi_{model}^2(\rho)}{\partial \rho_i} &= \frac{\partial \ln(|\mathbf{V}(\rho)|)}{\partial \rho_i} + \frac{\partial (\mathbf{R}(\rho)^T \mathbf{V}^{-1}(\rho) \mathbf{R}(\rho))}{\partial \rho_i} \end{aligned}$$

Nous allons utiliser les égalités suivantes sur la dérivée du logarithme du déterminant d'une matrice et sur la dérivée de l'inverse d'une matrice. Soit \mathbf{X} une matrice inversible alors :

$$\frac{\partial \ln(|\mathbf{X}|)}{\partial x} = \text{Tr} \left(\mathbf{X}^{-1} \frac{\partial \mathbf{X}}{\partial x} \right) \quad (6.39)$$

$$\frac{\partial \mathbf{X}^{-1}}{\partial x} = -\mathbf{X}^{-1} \frac{\partial \mathbf{X}}{\partial x} \mathbf{X}^{-1} \quad (6.40)$$

En notant, la matrice de $\mathbf{J}_{V,i} = \frac{\partial \mathbf{V}(\rho)}{\partial \rho_i}$ et $\mathbf{J}_i = \frac{\partial \mathbf{M}(\rho)}{\partial \rho_i}$, nous obtenons, sachant que : $\frac{\partial \mathbf{R}(\rho)}{\partial \rho_i} = \frac{\partial (\mathbf{Y} - \mathbf{M}(\rho))}{\partial \rho_i} = -\mathbf{J}^i$:

$$\frac{\partial \chi_{model}^2(\rho)}{\partial \rho_i} = \text{Tr} (\mathbf{V}^{-1} \mathbf{J}_{V,i}) - 2\mathbf{J}_i^T \mathbf{V}^{-1} \mathbf{R} - \mathbf{R}^T \mathbf{V}^{-1} \mathbf{J}_{V,i} \mathbf{V}^{-1} \mathbf{R}$$

Quant aux dérivées secondes, on utilise le fait que la dérivée de la trace d'une matrice est égale à la trace de la dérivée, pour écrire :

$$\begin{aligned} \frac{\partial^2 \chi_{model}^2(\rho)}{\partial \rho_j \partial \rho_i} &= \frac{\partial (\text{Tr} (\mathbf{V}^{-1} \mathbf{J}_{V,i}))}{\partial \rho_j} - \frac{\partial (2\mathbf{J}_i^T \mathbf{V}^{-1} \mathbf{R})}{\partial \rho_j} - \frac{\partial (\mathbf{R}^T \mathbf{V}^{-1} \mathbf{J}_{V,i} \mathbf{V}^{-1} \mathbf{R})}{\partial \rho_j} \\ &= \text{Tr} \left(-\mathbf{V}^{-1} \mathbf{J}_{V,j} \mathbf{V}^{-1} \mathbf{J}_{V,i} + \mathbf{V}^{-1} \frac{\partial \mathbf{J}_{V,i}}{\partial \rho_j} \right) - 2 \frac{\partial \mathbf{J}_{V,i}}{\partial \rho_j} \mathbf{V}^{-1} \mathbf{R} + 2\mathbf{J}_i^T \mathbf{V}^{-1} \mathbf{J}_{V,j} \mathbf{V}^{-1} \mathbf{R} + 2\mathbf{J}_i^T \mathbf{V}^{-1} \mathbf{J}_j^T \\ &\quad + 2\mathbf{J}_i^T \mathbf{V}^{-1} \mathbf{J}_{V,i} \mathbf{V}^{-1} \mathbf{R} + 2\mathbf{R}^T \mathbf{V}^{-1} \mathbf{J}_{V,j} \mathbf{V}^{-1} \mathbf{J}_{V,i} \mathbf{V}^{-1} \mathbf{R} - \mathbf{R}^T \mathbf{V}^{-1} \frac{\partial \mathbf{J}_{V,i}}{\partial \rho_j} \mathbf{V}^{-1} \mathbf{R} \end{aligned}$$

En considérant uniquement les dérivées premières du modèle, on supprime le second terme de la trace et le second et le dernier terme, pour obtenir finalement :

$$\frac{\partial^2 \chi_{model}^2(\rho)}{\partial \rho_j \partial \rho_i} = \text{Tr}(-\mathbf{V}^{-1} \mathbf{J}_{V,j} \mathbf{V}^{-1} \mathbf{J}_{V,i}) + 2\mathbf{J}_i^T \mathbf{V}^{-1} \mathbf{J}_{V,j} \mathbf{V}^{-1} \mathbf{R} \quad (6.41)$$

$$+ 2\mathbf{J}_i^T \mathbf{V}^{-1} \mathbf{J}_j^T + 2\mathbf{J}_i^T \mathbf{V}^{-1} \mathbf{J}_{V,i} \mathbf{V}^{-1} \mathbf{R} + 2\mathbf{R}^T \mathbf{V}^{-1} \mathbf{J}_{V,j} \mathbf{V}^{-1} \mathbf{J}_{V,i} \mathbf{V}^{-1} \mathbf{R} \quad (6.42)$$

Si l'on considère, comme dans l'algorithme de Fisher, seulement l'espérance du Hessien, alors puisque la moyenne des résidus est nulle, on se retrouve avec un formule plus simple :

$$\begin{aligned} \mathbf{H} = (\mathbf{H})_{ij} &= \left(\frac{\partial^2 \chi_{model}^2(\rho)}{\partial \rho_j \partial \rho_i} \right)_{ij} \\ &= (\text{Tr}(-\mathbf{V}^{-1} \mathbf{J}_{V,j} \mathbf{V}^{-1} \mathbf{J}_{V,i}) + 2\mathbf{J}_i^T \mathbf{V}^{-1} \mathbf{J}_j^T)_{ij} \end{aligned} \quad (6.43)$$

6.3.2 Dérivé du χ_{cons}^2

En repartant de l'équation 6.20, on peut calculer le gradient et le hessien :

$$\begin{aligned} \frac{\partial \chi_{cons}^2}{\partial \rho_i} &= \frac{\partial (\mu_{pen} \mathbf{C}(\rho)^T \mathbf{C}(\rho))}{\partial \rho_i} \\ &= 2\mu_{pen} \frac{\partial \mathbf{C}(\rho)^T}{\partial \rho_i} \mathbf{C}(\rho) \end{aligned} \quad (6.44)$$

Pour le hessien :

$$\frac{\partial^2 \chi_{cons}^2}{\partial \rho_j \partial \rho_i} = 2\mu_{pen} \left(\frac{\partial^2 \mathbf{C}(\rho)^T}{\partial \rho_j \partial \rho_i} \mathbf{C}(\rho) + \frac{\partial \mathbf{C}(\rho)^T}{\partial \rho_i} \frac{\partial \mathbf{C}(\rho)}{\partial \rho_j} \right) \quad (6.45)$$

Pour les six contraintes linéaires le calcul des dérivées est trivial. Les contraintes s'exprimant : $\mathbf{C}(\rho) = \alpha - \mathbf{H}_{pen} \rho$, les dérivées secondes sont nulles et nous obtenons :

$$\frac{\partial \chi_{cons}^2}{\partial \rho_i} = 2\mu_{pen} \mathbf{H}^T (\alpha - \mathbf{H}_{pen} \beta) \quad \& \quad \frac{\partial^2 \chi_{cons}^2}{\partial \rho_j \partial \rho_i} = 2\mu_{pen} \mathbf{H}^T \mathbf{H} \quad (6.46)$$

En revanche, la contrainte sur la variance des X_1 n'est pas linéaire. Durant la minimisation, nous aurons à calculer les dérivés première et seconde de cette contrainte en fonction des X_1 . C'est pourquoi l'algèbre est dérivée maintenant. On note la partie de l'ajustement par moindre carré correspondant à cette contrainte : $\chi_{var X_1}^2 = \mu_{pen}^1 \mathbf{C}^1(X_1)^T \mathbf{C}^1(X_1)$. On peut calculer le gradient de cette fonction :

$$\frac{\partial \chi_{var X_1}^2}{\partial \mathbf{X}_1^i} = 2\mu_{pen} \mathbf{C}^1(X_1)^T \frac{\partial \mathbf{C}^1(X_1)}{\partial \mathbf{X}_1^i}$$

et le hessien :

$$\frac{\partial^2 \chi_{var X_1}^2}{\partial \mathbf{X}_1^j \partial \mathbf{X}_1^i} = 2\mu_{pen} \left[\mathbf{C}^1(X_1)^T \frac{\partial^2 \mathbf{C}^1(X_1)}{\partial \mathbf{X}_1^j \partial \mathbf{X}_1^i} + \frac{\partial \mathbf{C}^1(X_1)^T}{\partial \mathbf{X}_1^j} \frac{\partial \mathbf{C}^1(X_1)}{\partial \mathbf{X}_1^i} \right]$$

On écrivant la contrainte sous forme matricielle, les dérivées s'obtiennent directement :

$$\mathbf{C}^1(X_1) = \frac{1}{N} \mathbf{X}_1^T \mathbf{X}_1 - 1 \quad \& \quad \frac{\partial \mathbf{C}^1(X_1)}{\partial \mathbf{X}_1^i} = \frac{2}{N^2} \mathbf{X}_1^T \quad \& \quad \frac{\partial^2 \mathbf{C}^1(X_1)}{\partial \mathbf{X}_1^j \partial \mathbf{X}_1^i} = \frac{2}{N^2} \delta_{ij} \quad (6.47)$$

6.3.3 Dérivé du χ_{Reg}^2

Les dérivées du terme de régularisation sont extrêmement simples. En repartant de l'équation 6.21, nous obtenons :

$$\nabla \chi_{reg}^2 = 2\mu_{reg}\rho^T \mathbf{P} \quad \mathbf{H}_{\chi_{reg}^2} = 2\mu_{reg}\mathbf{P} \quad (6.48)$$

6.3.4 Dérivé du χ_{calib}^2

On obtient aussi facilement le gradient et le hessien du terme de propagation des incertitudes de calibration :

$$\nabla \chi_{calib}^2 = 2\rho^T \mathbf{V}_\eta^{-1} \quad \mathbf{H}_{\chi_{calib}^2} = 2\mathbf{V}_\eta^{-1} \quad (6.49)$$

6.3.5 Dérivé du χ_{CS}^2

En repartant de l'équation 6.37, et en distinguant pour plus de clarté les paramètres κ et σ_κ , on obtient pour le gradient par rapport aux paramètres κ :

$$\begin{aligned} \frac{\partial \chi_{CS}^2}{\partial \kappa_i} &= \frac{\partial (\ln |\mathbf{V}_\kappa(\sigma_\kappa)| + \kappa^T \mathbf{V}_\kappa^{-1}(\sigma_\kappa) \kappa)}{\partial \kappa_i} \\ &= 2\mathbf{V}_\kappa^{-1} \kappa \end{aligned} \quad (6.50)$$

Pour le gradient par rapport aux σ_κ , on utilise les équations 6.39 et 6.40, pour obtenir :

$$\begin{aligned} \frac{\partial \chi_{CS}^2}{\partial \sigma_{\kappa,i}} &= \frac{\partial (\ln |\mathbf{V}_\kappa(\sigma_\kappa)| + \kappa^T \mathbf{V}_\kappa^{-1}(\sigma_\kappa) \kappa)}{\partial \sigma_{\kappa,i}} \\ &= \text{Tr} \left(\mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,i}} \right) - \kappa^T \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,i}} \mathbf{V}_\kappa^{-1} \kappa \end{aligned} \quad (6.51)$$

Il y a ainsi trois contributions aux hessiens, les dérivées de chaque groupe de paramètres et les dérivées croisées :

$$\frac{\partial^2 \chi_{CS}^2}{\partial \kappa_j \partial \kappa_i} = 2\mathbf{V}_\kappa^{-1} \quad \frac{\partial^2 \chi_{CS}^2}{\partial \kappa_j \partial \sigma_{\kappa,i}} = -2\mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,i}} \mathbf{V}_\kappa^{-1} \kappa \quad (6.52)$$

$$\begin{aligned} \frac{\partial^2 \chi_{CS}^2}{\partial \sigma_{\kappa,j} \partial \sigma_{\kappa,i}} &= \text{Tr} \left(-\mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,j}} \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,i}} + \mathbf{V}_\kappa^{-1} \frac{\partial^2 \mathbf{V}_\kappa}{\partial \sigma_{\kappa,j} \partial \sigma_{\kappa,i}} \right) \\ &\quad + \kappa^T \mathbf{V}_\kappa^{-1} \left[2 \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,j}} \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,i}} - \frac{\partial^2 \mathbf{V}_\kappa}{\partial \sigma_{\kappa,j} \partial \sigma_{\kappa,i}} \right] \mathbf{V}_\kappa^{-1} \kappa \end{aligned} \quad (6.53)$$

L'espérance de la dérivée croisée entre les paramètres κ et σ_κ étant nulle nous ne considérons pas ce terme. La dérivé seconde par rapport aux coefficients du polynôme est nulle, donc :

$$\frac{\partial^2 \chi_{CS}^2}{\partial \sigma_{\kappa,j} \partial \sigma_{\kappa,i}} = -\text{Tr} \left(\mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,j}} \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,i}} \right) + 2\kappa^T \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,i}} \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,j}} \mathbf{V}_\kappa^{-1} \kappa \quad (6.54)$$

6.3.6 L'équation finale

L'équation normale finale est donc la combinaison des dérivées premières et secondes de toutes ces contributions. D'abord le gradient peut se réécrire sous la forme matricielle suivante, en remplaçant les paramètres ρ par les différents paramètres :

$$\nabla\chi^2 = \left(\begin{array}{c} \text{Tr} \left(\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \beta} \right) - 2 \frac{\partial \mathbf{M}}{\partial \beta}^T \mathbf{V}^{-1} \mathbf{R} - \mathbf{R}^T \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \beta} \mathbf{V}^{-1} \mathbf{R} \\ + 2 \mu_{pen} \frac{\partial \mathbf{C}(\beta)}{\partial \beta}^T \mathbf{C}(\beta) \\ + 2 \mu_{reg} \beta^T \mathbf{P} + 2 \beta^T \mathbf{V}_\eta^{-1} + 2 \kappa^T \mathbf{V}_\kappa^{-1} \\ \text{-----} \\ \text{Tr} \left(\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \gamma} \right) - \mathbf{R}^T \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \gamma} \mathbf{V}^{-1} \mathbf{R} \\ \text{-----} \\ \text{Tr} \left(\mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_\kappa} \right) - \kappa^T \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_{\kappa,i}} \mathbf{V}_\kappa^{-1} \kappa \end{array} \right) \begin{array}{l} \beta \\ \gamma \\ \sigma_\kappa \end{array} \quad (6.55)$$

$$\mathbf{H} = \left(\begin{array}{ccc} \overbrace{\text{Tr} \left(-\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \beta} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \beta} \right) + 2 \frac{\partial \mathbf{M}(\beta)}{\partial \beta}^T \mathbf{V}^{-1} \frac{\partial \mathbf{M}(\beta)}{\partial \beta}}^{\beta} & \overbrace{\dots}^{\gamma} & \overbrace{\dots}^{\sigma_\kappa} \\ + 2 \mu_{pen} \left(\frac{\partial^2 \mathbf{C}(\beta)}{\partial \beta \partial \beta}^T \mathbf{C}(\beta) + \frac{\partial \mathbf{C}(\beta)}{\partial \beta}^T \frac{\partial \mathbf{C}(\beta)}{\partial \beta} \right) & \text{Tr} \left(-\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \gamma} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \beta} \right) & 0 \\ + 2 \mathbf{V}_\eta^{-1} + 2 \mathbf{V}_\kappa^{-1} + 2 \mu_{reg} \mathbf{P} & \dots & \dots \\ \text{-----} \\ \text{Tr} \left(-\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \beta} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \gamma} \right) & \text{Tr} \left(-\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \gamma} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \gamma} \right) & 0 \\ \text{-----} \\ 0 & 0 & \text{Tr} \left(-\mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_\kappa} \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \sigma_\kappa} \right) \\ \dots & \dots & + 2 \kappa^T \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \rho_j} \mathbf{V}_\kappa^{-1} \frac{\partial \mathbf{V}_\kappa}{\partial \rho_i} \mathbf{V}_\kappa^{-1} \kappa \end{array} \right) \begin{array}{l} \beta \\ \gamma \\ \sigma_\kappa \end{array} \quad (6.56)$$

Pour chaque itération, on cherche à calculer l'incrément $\delta\beta$ défini à l'équation 6.17 :

$$\delta\beta = -\mathbf{H}^{-1} \nabla\chi^2 \quad (6.57)$$

En pratique, si on considère 1000 SNe, soit 18332 paramètres, la matrice \mathbf{H} sera de 18332×18332 paramètres. Inverser une matrice de cette taille prend beaucoup de temps. Par chance, la hessienne n'est pas dense, le modèle des différentes observations est peu corrélé d'une SN à l'autre. On peut donc exploiter les avantages des matrices creuses, dites *sparse*, définies dans le module *scipy* et la fonction *cholesky* du module *scikit-sparse*. Cette fonction permet d'inverser rapidement des matrices creuses sous certaines conditions de densité. La matrice est composée à

plus de 75% de valeurs nulles. Plus le nombre de SNe dans le lot d'entraînement augmente moins elle sera remplie.

L'inversion de la matrice se fait en utilisant la méthode de Cholesky. Elle permet de résoudre efficacement les systèmes linéaires de forme $\mathbf{Ax} = \mathbf{y}$, avec \mathbf{A} définie positive en exploitant ses symétries. L'algorithme se décompose en trois étapes. La complexité d'un algorithme quantifie son efficacité, c'est une fonction du nombre de données noté n . Si \mathbf{A} est une matrice symétrique définie positive alors il existe une unique matrice triangulaire inférieure \mathbf{L} tel que : $\mathbf{A} = \mathbf{LL}^T$, sa décomposition de Cholesky. La détermination de cette matrice est de complexité cubique, $O(n^3)$. L'étape de "descente" correspond à la résolution du système $\mathbf{Ly} = \mathbf{b}$, suivi de la "remontée" résolvant $\mathbf{L}^T \mathbf{x} = \mathbf{y}$. Ces deux étapes sont de complexité quadratique $O(n^2)$. La complexité totale, cubique permet de résoudre rapidement des systèmes de plusieurs milliers de paramètres.

6.4 Estimation des incertitudes sur les paramètres

Puisque nous ajustons tous les paramètres simultanément, l'erreur sur les paramètres est un produit naturel de l'ajustement. En repartant de l'équation 6.3, on peut exprimer la quantité d'information dans l'observation de \mathbf{Y} sur le paramètre β , noté $\mathbf{I}_Y(\beta)$:

$$\mathbf{I}_Y(\beta) = \mathbb{E} \left[\left(\frac{\partial \ln L(\mathbf{Y}|\beta)}{\partial \beta} \right)^2 \right] \quad (6.58)$$

Si l'espace des paramètres des β ne dépend pas de leur valeurs et que la fonction L est assez régulière pour permuter, dérivée, dérivée seconde et intégrale (comme dans notre cas), alors on peut ré-écrire l'équation 6.58 comme :

$$\mathbf{I} = [\mathbf{I}_X(\beta)]_{ij} = -\mathbb{E} \left[\frac{\partial^2 \ln L(\mathbf{Y}|\beta)}{\partial \beta_i \partial \beta_j} \right] \quad (6.59)$$

\mathbf{I} est l'information de Fisher sur les paramètres θ extraites des observations \mathbf{X} . Dans notre cas, puisque nous utilisons un algorithme de score de Fisher, cette information est directement donné par la matrice hessienne \mathbf{H} (equation 6.56) :

$$\mathbf{I} = \mathbb{E} \left[-2 \frac{\partial^2 \ln L(\mathbf{X}|\beta)}{\partial \beta_i \partial \beta_j} \right] = \frac{\mathbf{H}}{2} \quad (6.60)$$

À l'issue de chaque minimisation sur l'ensemble des paramètres, la variance sur chaque paramètre est égale à la diagonale de l'inverse de la matrice de Fisher :

$$\mathbf{V}(\beta_i) = 2\mathbf{H}_{ii}^{-1} \quad (6.61)$$

6.5 Conclusion

Nous avons développé un code spécifique afin minimiser des log-vraisemblances construites à partir de modèles potentiellement non-linéaires avec optionnellement un modèle d'erreur éventuellement non-diagonal, des contraintes quadratiques et de la régularisation. La stratégie de minimisation est organisée autour d'un algorithme de type Gauss-Newton. L'interface du minimiseur est générique. Il peut donc être utilisé avec une grande classe de modèles alternatifs. Ainsi, de nouvelles paramétrisations sont possibles : développement de la PCA aux ordres supérieurs, ajout de nouveaux paramètres tels que la couleur de la galaxie hôte, ou une couleur tardive de la SN...

Le minimiseur permet de construire la matrice de Fisher à l'issue de l'ajustement. Les incertitudes sont propagées de manière exacte. La matrice de covariance des paramètres propres aux SNe est un produit direct de la minimisation. Aucune opération supplémentaire n'est requise.

Une part importante du travail s'est portée sur l'accélération de cet outil, permettant d'ajuster un modèle de plusieurs dizaines de milliers de paramètres sur plusieurs milliers de SNe en seulement quelques heures. L'entraînement sur des lots réalistes comme ceux de SALT2.4 et SALT3 se réalise en respectivement 30 minutes et une heure. Cette procédure rapide nous permet de tester rapidement sur des simulations les biais éventuels du modèle. Pour développer le modèle et le minimiseur, puis pour tester les stratégies de minimisation, et enfin pour quantifier les biais éventuels affectant les paramètres reconstruits, ainsi que l'efficacité de l'estimateur, nous nous sommes appuyés sur des échantillons d'entraînement simulés. La génération de ces données fait l'objet du prochain chapitre.

Simulations

Les outils décrits dans la section précédente ont été développés sur des simulations. Nous avons développé un code permettant de générer rapidement des lots d'entraînement photométriques et spectroscopiques, en utilisant des modèles des principaux instruments utilisés par la communauté (VRO, DECam, MegaCam, SDSS, Subaru, SWOPE...)

Dans un premier temps, nous avons testé nos codes d'ajustement avec des lots de données "parfaits" : cadence élevée, courbes de lumière bien échantillonnées dans 4 à 5 bandes par SN, un ou plusieurs spectres par SN et un signal sur bruit indépendant du *redshift*. Le code de simulation part d'une cosmologie fiducielle, des modèles de télescopes et des surfaces SALT2.4 pour produire des lots de données de taille arbitraire.

Nous avons ensuite développé de quoi générer des lots de données plus réalistes, avec un signal sur bruit dépendant du *redshift*, en utilisant un modèle du télescope Vera Rubin. Ces simulations idéales dites "LSST-like", nous ont permis de tester les performances de notre code, en étudiant les biais de reconstruction et l'efficacité de l'estimateur de distance.

Enfin, nous nous sommes dotés de la possibilité de reproduire un nombre arbitraire de réalisations des lots JLA et K21, qui ont servi à entraîner respectivement SALT2.4 et SALT3. Ces réalisations sont réalistes (mêmes objets, même sampling, même signal sur bruit) et permettent d'étudier les biais éventuels du modèle sur les échantillons actuels. Les lots de données de type LSST permettent de passer continûment de l'état de l'art actuel aux lots d'entraînement futurs.

Contents

7.1	Simulation de différents lot d'entraînement	160
7.1.1	Simulations idéales	160
7.1.2	Lots de type LSST	162
7.1.3	Simulation lots JLA et K21	163
7.2	Flux	171
7.3	Génération du bruit ajouté	172
7.3.1	Exemple SNe de JLA	173
7.4	Conclusion	174

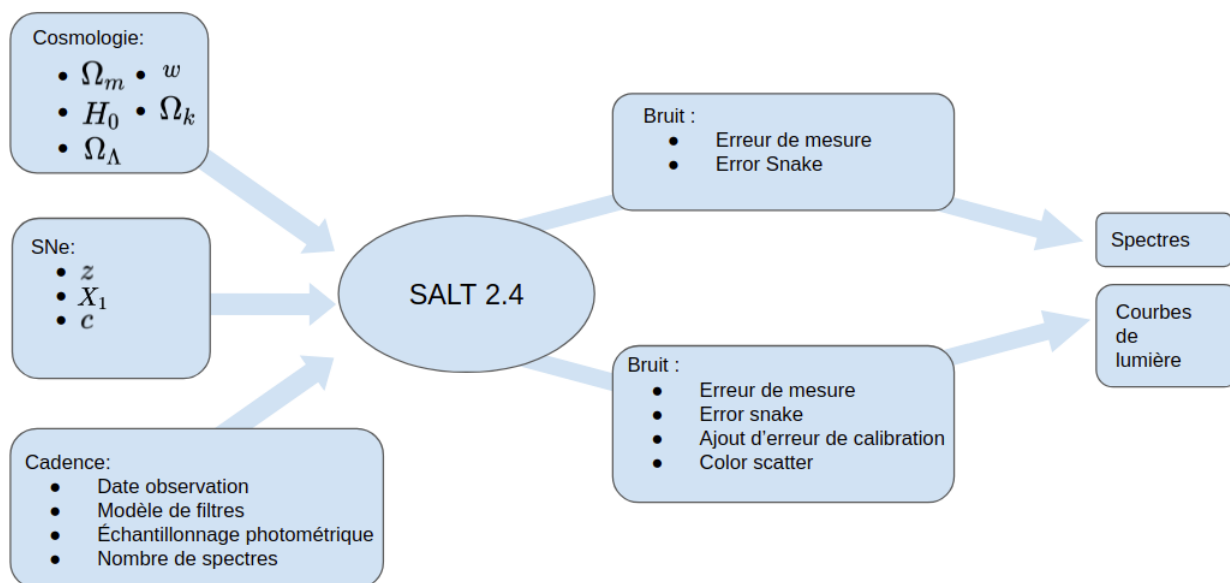


FIGURE 7.1 – Schéma du module de simulation. Il est composé de trois étapes. Lors de la première : on crée les dates et les caractéristiques des observations et les paramètres propres à chaque SN, en supposant une cosmologie pour les simulations idéales et pour les lots réalistes, les observations sont recopiées et les paramètres d’un ajustement de SALT2.4 sont utilisés. La seconde étape est la modélisation du flux, en utilisant le modèle SALT2.4. La dernière étape ajoute du bruit aux spectres et courbes de lumière.

7.1 Simulation de différents lot d’entraînement

Nous présentons figure 7.1 l’organisation de notre pipeline de simulation. Il est divisé en trois segments. Le premier segment consiste en la génération de la cosmologie fiducielle, du jeu de SNe et de la cadence d’observation, pour les simulations réalistes. Concernant les émulations des lots de données JLA et K21, ces trois ingrédients sont inutiles : le lot de données et la cadence sont imposés.

Le segment central est la génération des courbes de lumière et des spectres. Elle est réalisée en utilisant notre modèle, initialisé avec les surfaces et la loi de couleur de SALT2.4 Pour les courbes de lumière, le modèle est intégré dans les modèles d’instruments correspondant à chaque point de la cadence.

Enfin, le dernier segment consiste à générer puis appliquer aux données le bruit de mesure, ainsi que la variance additionnelle (calibration, *error snake*, *color scatter*). Pour les données du relevé LSST, l’évaluation du bruit de mesure à ajouter aux données est réalisée grâce à un *exposure time calculator* (ETC). Pour l’émulation des données JLA et K21, nous utilisons directement les incertitudes publiées, associées à chaque point de mesure afin de reproduire exactement le signal sur bruit des données originales.

7.1.1 Simulations idéales

Ce premier jeu de simulations est important car il permet un contrôle fin de toutes les grandeurs intervenant dans la simulation. Nous choisissons le nombre de supernovæ voulu, les distributions des *redshift*, X_1 , c et $t_{max}^{B^*}$ des objets. Pouvoir varier le nombre de SNe, le nombre de spectres et de

courbes de lumière par objet, ainsi que la cadence d'observations ou la résolution spectroscopique permet de délimiter le domaine de fonctionnement du modèle et de tester sa rapidité. Pour les deux types de simulations idéales, ce processus est le même et les grandeurs intervenant dans la première partie du *pipeline* sont les mêmes, mise à part le modèles des instruments (cf. Figure 7.1).

Paramètres des SNe Pour un nombre de supernovæ fixé, on tire uniformément¹ des *redshifts* compris par défaut entre $z = 0.001$ et $z = 0.8$. Les lots d'entraînement sont tronqués en *redshift* par rapport aux lots cosmologiques dans le but d'avoir un lot d'entraînement complet, sans biais de sélection.

Une fois les *redshifts* obtenus, on calcule la distance de luminosité définie équation 1.16. Notre cosmologie fiducielle est la cosmologie Flat- Λ CDM suivante :

- $H_0 = 70 \text{ km.s}^{-1}.\text{Mpc}^{-1}$
- $w = -1$
- $\Omega_k = 0$.
- $\Omega_m = 0.25$
- $\Omega_\Lambda = 0.75$
- $\Omega_\gamma = 5 * 10^{-5}$

Une fois les distances connues, on peut simuler les paramètres propres des SNe. On génère d'abord les valeurs de X_1 et c selon les deux lois normales suivantes :

$$X_1 \sim \mathcal{N}(0, 1.) \quad c \sim \mathcal{N}(0, 0.1)$$

Les distributions de X_1 et c ont toutes deux des queues non gaussiennes bien décrites par les observations que nous omettons pour des raisons de simplifications. On calcule ensuite X_0 selon la formule suivante :

$$X_0 = \frac{10^{-4}}{d_L(z)^2} 10^{-0.4(\tilde{M}_{B^*} - \alpha X_1 + \beta c)} \quad (7.1)$$

Le facteur 10^{-4} correspond au passage du carré de la distance de luminosité de kpc^2 en pc^2 . Pour les paramètres de standardisation, nous prenons ceux déterminés par l'analyse JLA :

- $\tilde{M}_{B^*} = -19.0906$
- $\alpha = 0.13$
- $\beta = 3$

Organisation du relevé

Pour les lots de données idéaux, la stratégie de prise de données est totalement configurable. Le nombre de courbes de lumière, d'observations par courbe de lumière ou de spectres par SN est laissé au choix de l'observateur/simulateur. Il est par ailleurs possible de générer soit plusieurs lots de SNe, suivis par des programmes différents, afin d'émuler les échantillons hétérogènes actuels, soit un lot obtenu par un relevé unique avec une couverture en *redshift* assez large pour définir un lot d'entraînement utilisable pour les études cosmologiques.

Filtres Dans ce premier jeu de simulations idéales, chaque SN est observée dans 4 bandes. L'instrument utilisé dépend du *redshift* de la SN :

- si $z < 0.05$: par B, V, r, i du SWOPE
- si $0.05 < z < 0.2$: par g, r, i, z du SDSS
- si $0.2 < z$: par g, r, i, z de MEGACAM

1. Cette distribution n'a aucun sens physique.

Cadence Par défaut, les courbes de lumières des lots idéaux sont composées de 20 points photométriques. Les phases sont tirées selon une loi uniforme entre -15 et $+35$ jours. Les dates d’observation dans le référentiel de l’observateur sont inférées selon la formule suivante :

$$date = phase(1 + z) + t_{max}^{B^*} \quad (7.2)$$

Spectres Par défaut, un spectre unique est simulé par SN. Le choix d’un spectre par SN est justifié par la figure 5.8, sur laquelle on voit que les lots d’entraînement utilisés pour SALT2.4 et SALT3 ont en moyenne un spectre par SN. La résolution spectrale est fixe et égale à 10 \AA . Les spectres sont observés entre λ_{bleu} et λ_{rouge} , respectivement tirés entre 200 et 300 nm et 800 et 900 nm. Les dates d’observation des spectres sont tirées selon une loi de probabilité extraite des dates observées des données de K21. Pour chaque spectre, on reconstitue la phase d’observation. L’histogramme des phases est représenté sur la figure 7.2. Cinq distributions statistiques ont été ajustées par moindre carré. La distribution décrivant le mieux les données est une distribution de

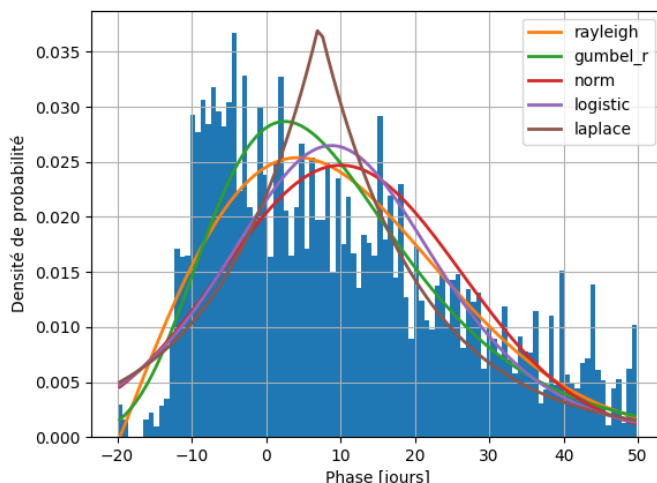


FIGURE 7.2 – Distribution des phases spectrales du lot d’entraînement de K21. Les cinq courbes représentent différentes fonctions de densité ajustées sur l’histogramme. La courbe verte correspondant à un ajustement minimal, représente une distribution de Gumbel dont les paramètres sont $\mu = 2.5$ et $\beta = 12.8$.

Gumbel, dont la densité de probabilité est définie par formule suivante :

$$f(x) = \frac{e^{-e^{-\frac{x-\mu}{\beta}}} e^{-\frac{x-\mu}{\beta}}}{\beta} \quad (7.3)$$

À l’issue de l’ajustement $\mu = -2.01$ et $\beta = 8.27$. Les phases d’observations des spectres sont donc tirées selon cette loi.

7.1.2 Lots de type LSST

Les lots de données idéaux décrits dans la section précédente se sont révélés très utiles dans les phases de développement et de test du code. En revanche, ils manquent de réalisme : les SNe

sont réparties uniformément en *redshift*, la qualité des mesures de flux ne dépend pas de z , enfin, la qualité du suivi est bien meilleure que ce qui sera possible de réaliser avec les moyens actuels.

Nous avons donc développé, dans un second temps, de quoi générer des lots de données plus réalistes, par exemple du type de celui qui sera produit par LSST. Nous remplaçons le modèle des instruments par celui du télescope Vera Rubin pour simuler un suivi LSST. Les dates d'observations, la cosmologie et les grandeurs des différentes supernovæ observées sont créés de la même manière. Le suivi est simulé pour reproduire les deux composantes principales du relevé LSST : le "Wide Fast Deep" (WFD) (visites de 30-s tous les 3-4 heures), pour des *redshifts* inférieurs à 0.4 et "Deep Drilling Fields" (DDF, visites de 600-s tous les 2-3 jours) au delà de cette limite.

Par défaut, nous supposons que la stratégie d'observation du relevé est suffisante pour produire pour chaque SN, des courbes de lumière dans les cinq bandes d'observation u , g , r , i et z (cf. Figure 3.10). Chaque courbe de lumière est échantillonnée avec 20 visites. Il s'agit de simulations idéales pour accompagner le développement du code, même le suivi LSST ne pourra pas fournir une telle cadence. Elle peut être dégradée continûment pour passer des simulations idéales, créés pour le développement, aux simulations réalistes des différents lot d'entraînement.

7.1.3 Simulation lots JLA et K21

Enfin, comme il était essentiel de s'assurer que le modèle pouvait être entraîné sur les lots de données actuels, nous avons aussi simulé, de la manière la plus réaliste possible, les lots d'entraînement JLA et K21. Nous repartons des objets composant chacun de ces lots d'entraînement : nous reprenons leurs *redshifts*, leurs paramètres SALT2 (X_0 , X_1 , c , t_{max}) ainsi les dates d'observation photométriques et spectroscopiques. À partir de ces informations, et du modèle SALT2.4, nous générons des courbes de lumière et des spectres qui sont ensuite bruités en utilisant les incertitudes publiées.

JLA comme K21 sont des lots composites, résultant de l'assemblage de plusieurs relevés. Les différents sous-échantillons photométriques de JLA et K21 sont désignés par le nom du relevé, par celui de l'instrument d'observation ou encore par le système de filtres utilisé. Comme leurs dénominations diffèrent entre les deux collaborations, nous les détaillons dans le tableau 7.1.

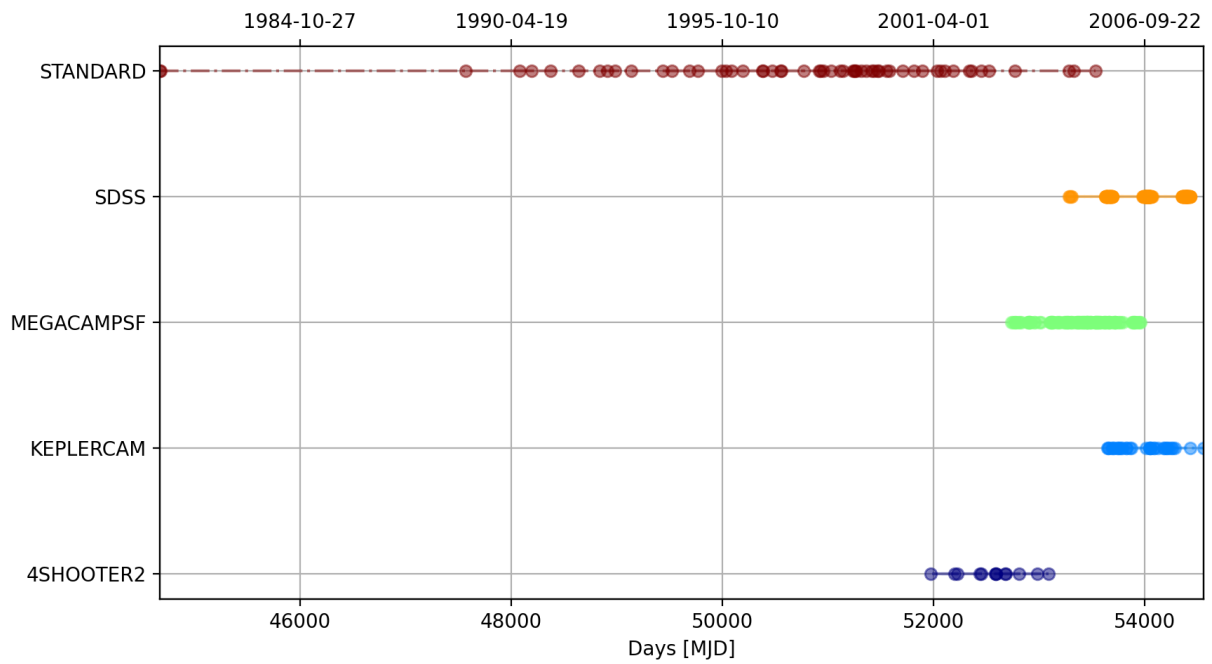
Relevé	JLA	K21
Relevés historiques	STANDARD	Bessell
CSP	-	CSP
CfA3	4SHOOTER2	CFA3S
CfA3	KEPLERCAM	CFA3K
CfA4	-	CFA41
SDSS	SDSS	SDSS
SNLS	MEGACAMPSF	SNLS
Pan-Starr	-	PS1
DES	-	DES

TABLE 7.1 – Code des noms des différents relevés utilisés dans les lots JLA et K21.

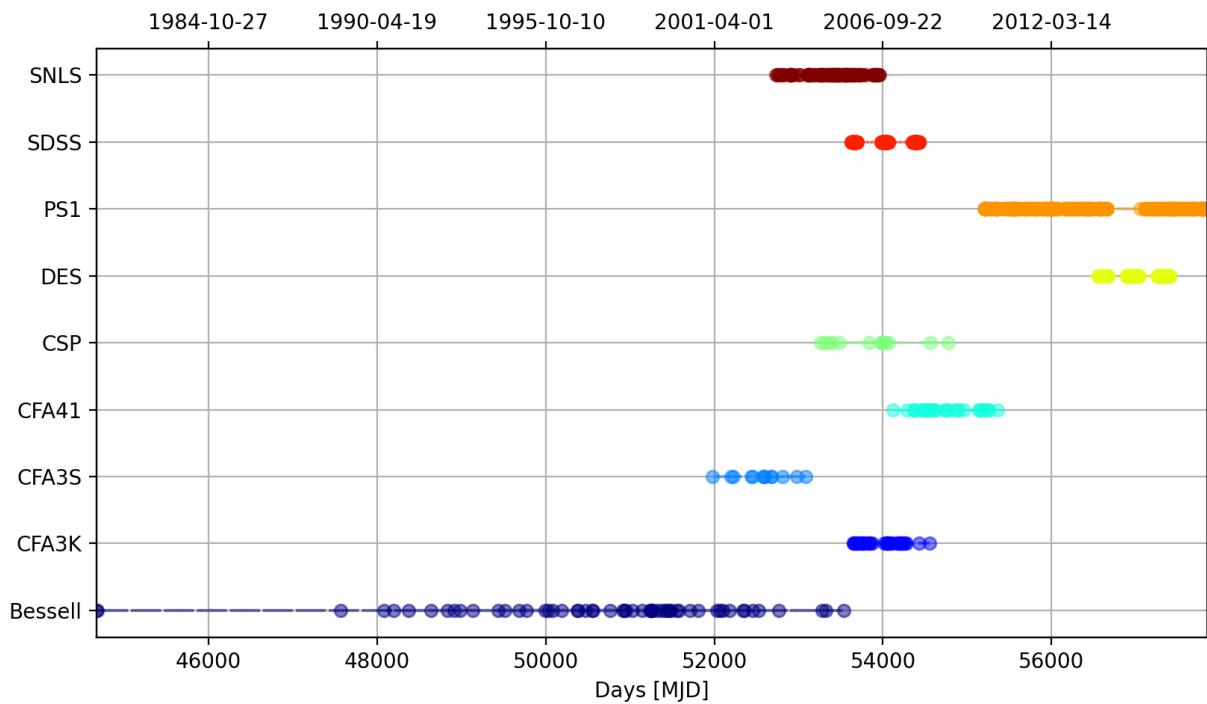
Le lot d'entraînement de K21, reprend les SNe de JLA avec l'ajout des données de l'échantillon *Foundation*, 153 SNe ont été ajoutées, dans la gamme de redshift $0.005 < z < 0.111$, auxquelles s'ajoutent 206 SNe du relevé DES, dont les *redshifts* sont compris entre $0.078 < z < 0.85$, et 266 SNe de PS1 ($0.026 < z < 0.63$).

Sur les figures 7.3a et 7.3b on peut voir les distributions de la date du maximum de luminosité

en bandes B^* des SNe composant les deux lots d'entraînement par relevé.



(a) Distribution des maxima de luminosité par relevé composant JLA.



(b) Distribution des maxima de luminosité par relevé composant K21.

FIGURE 7.3 – Distribution des maxima de luminosité de chaque relevé composant JLA (haut) et K21 (bas).

Le nombres de spectres et de courbes de lumières utilisés pour les simulations idéales, ainsi que la qualité de l'échantillonnage sont extrapolées à partir de ces deux lots d'entraînement.

Distributions des spectres On peut comparer les distributions des spectres pour les deux lots d'entraînement JLA et K21. Pour rappel, JLA contient 482 spectres appartenant à 83 SNe différentes. Quant à K21, il contient 1285 spectres de 271 SNe. On peut voir sur les figures 7.4 et 7.5, les distributions respectives du nombre de spectres par SNe, en fonction des relevés pour JLA et K21.

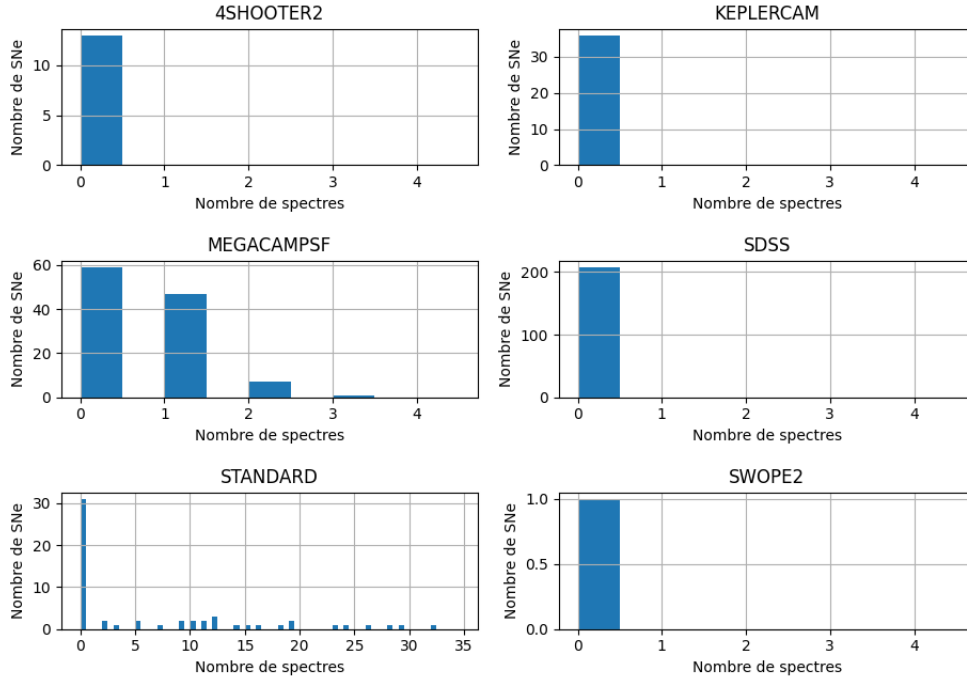


FIGURE 7.4 – Distribution du nombre de spectres par supernova composant chaque relevé du lot d'entraînement de JLA.

Pour JLA, les spectres proviennent des relevés historiques à bas *redshift* pour plus des trois quarts et de SNLS pour moins d'un quart. On peut voir sur la figure 7.5, que l'ajout de spectres du lot de K21 provient, en particulier, de DES et CfA. Nous avons bien un spectre par SN en moyenne, comme pour nos simulations, mais avec une distribution non homogène. Pour JLA et K21 les spectres sont essentiellement observés pour les SNe à bas *redshift*, avec quelques spectres pour SNLS.

Distributions des courbes de lumières Quant aux statistiques des courbes de lumière, le lot d'entraînement de JLA est composé de 1808 courbes de lumière appartenant à 425 SNe dont les *redshifts* vont de 0.01 à 0.8. K21 est composé de 3845 courbes de lumière appartenant à 1083 SNe dont les *redshifts* varient entre 0.0015 et 0.85. On peut voir sur les figures 7.6 et 7.7, les distributions du nombre de courbes de lumière par SN pour chaque relevé. Une supernova a été observée par KEPLERCAM, SDSS et SWOPE, et possède au total 15 bandes d'observations. Entre K21 et JLA, quelques SNe de CFA3 ont été ajoutées avec des observations dans plus de cinq bandes. Là encore les ajouts principaux proviennent de DES et PS1. Pour ces deux projets, les observations ont été réalisées dans quatre bandes. Ces deux figures justifient de choisir quatre ou cinq bandes d'observation pour nos simulations idéales.

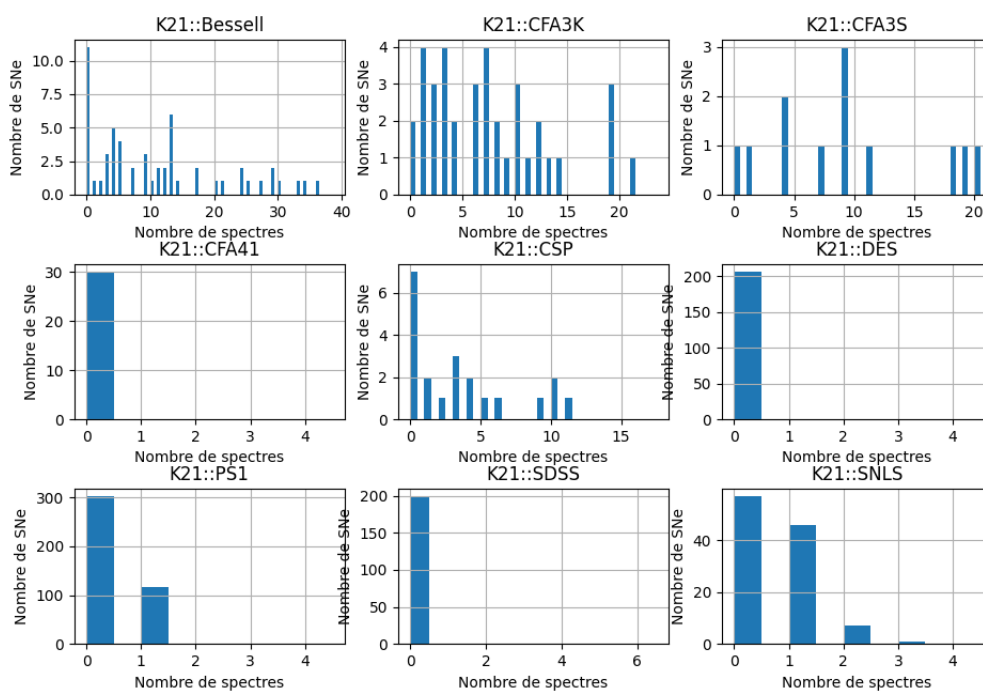


FIGURE 7.5 – Distribution du nombre de spectres par supernova composant chaque relevé du lot d’entraînement de K21.

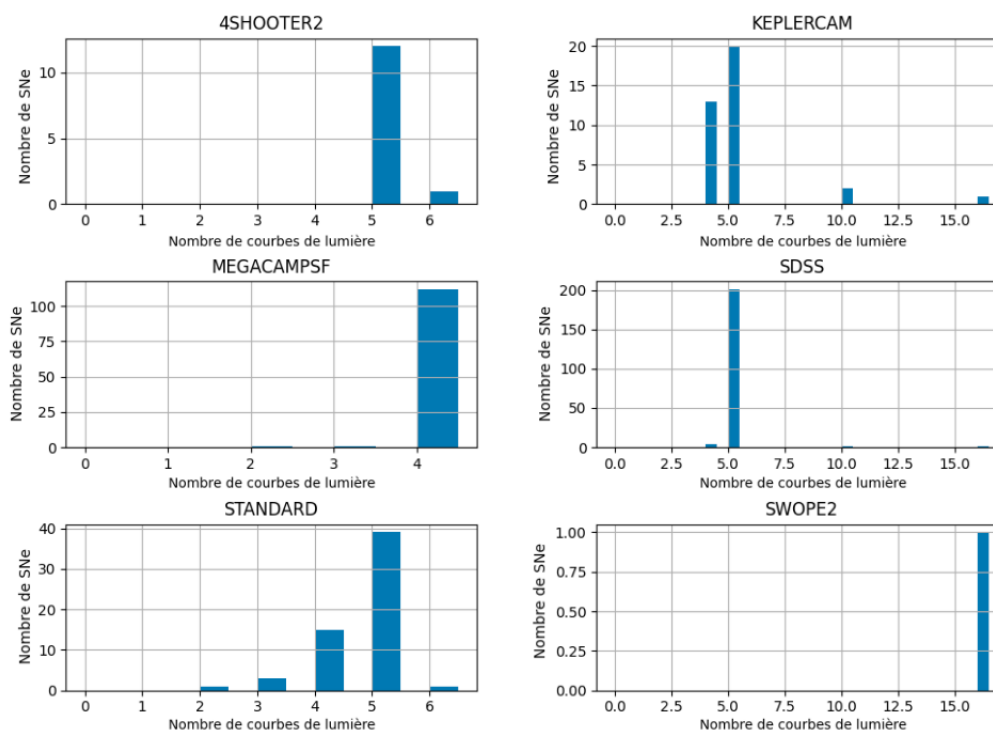


FIGURE 7.6 – Nombre de courbes de lumières par SNe en fonction des relevés composant JLA.

Les figures 7.7 et 7.8, montrent les échantillonnages des courbes de lumières pour les relevés de JLA et de K21 en fonction des filtres. Les courbes de lumières de SNLS et SDSS sont particulièrement bien échantillonnées, notamment grâce à la stratégie de recherche glissante mise en place

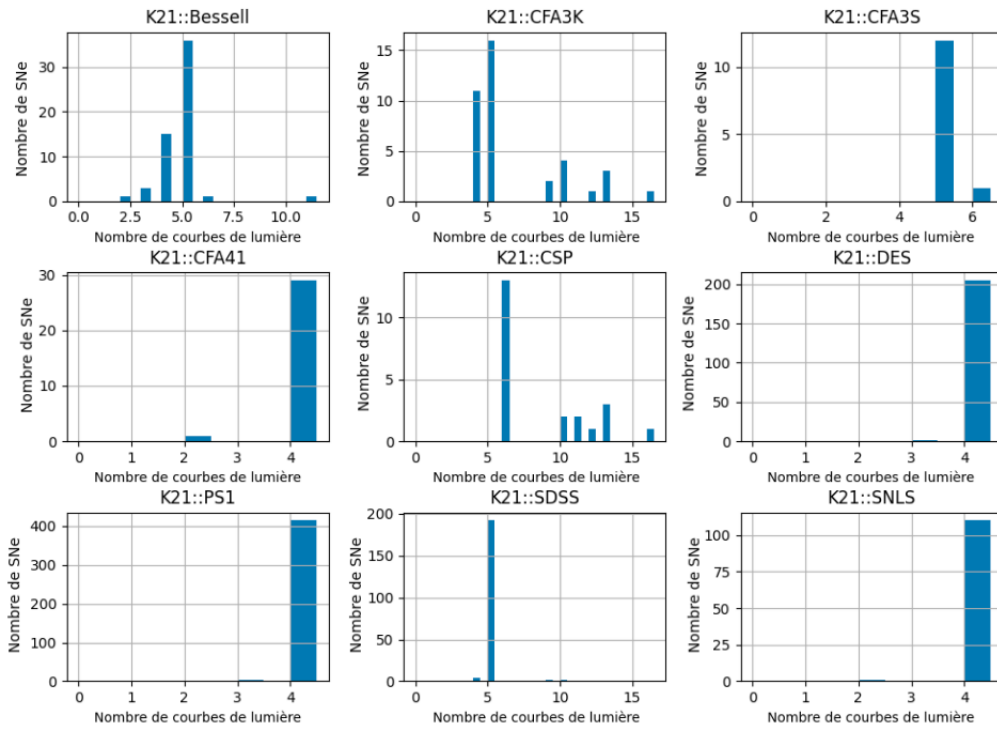


FIGURE 7.7 – Nombre de courbes de lumières par SNe en fonction des relevés composant K21.

par ces relevés. Les courbes de lumière de DES et PS1 ont une cadence inférieure à 20 points par courbe de lumière mais les distributions sont remarquablement homogènes. Enfin, le choix de vingt points par courbe de lumière pour nos simulations idéales est légèrement supérieur aux observations réalistes. On note enfin que pour nos simulations idéales, toutes les courbes de lumière ont des observations avant le maximum de luminosité, ce qui n'est pas le cas pour les lots réalistes.

Diagramme de couverture Puisque nous essayons d'extraire des surfaces à deux dimensions, M_0 et M_1 , en fonction de la longueur d'onde et de la phase, il est intéressant de représenter la quantité d'information disponible dans les lots d'entraînement dans un diagramme de couverture.

Pour ce faire, on définit une grille en phase et en longueur d'onde. Pour les courbes de lumière, on place les points de mesure pour chaque valeur de phase définie dans le référentiel de la SN et à la longueur d'onde moyenne du filtre corrigé du *redshift*. Pour les spectres, on compte un point dans la case correspondant à la phase et à la longueur d'onde, toutes deux corrigés du *redshift*. Tous les points ne sont pas comptés pour éviter les biais dus à la différence de résolution entre les spectres.

Le diagramme de couverture des données photométriques de JLA (respectivement K21) est visible sur la figure 7.10a (respectivement figure 7.11a), celui des données spectroscopiques est représenté figure 7.10b (respectivement figure 7.11b).

On peut y voir que le lot d'entraînement de K21 ajoute beaucoup de données photométriques sur toute la gamme de longueur d'onde. En revanche, la majorité des spectres est ajoutée dans la zone spectrale $3500 < \lambda < 7500 \text{ \AA}$ déjà bien contrainte par les données de JLA. Les régions infrarouge et ultraviolette doivent encore être mieux peuplées.

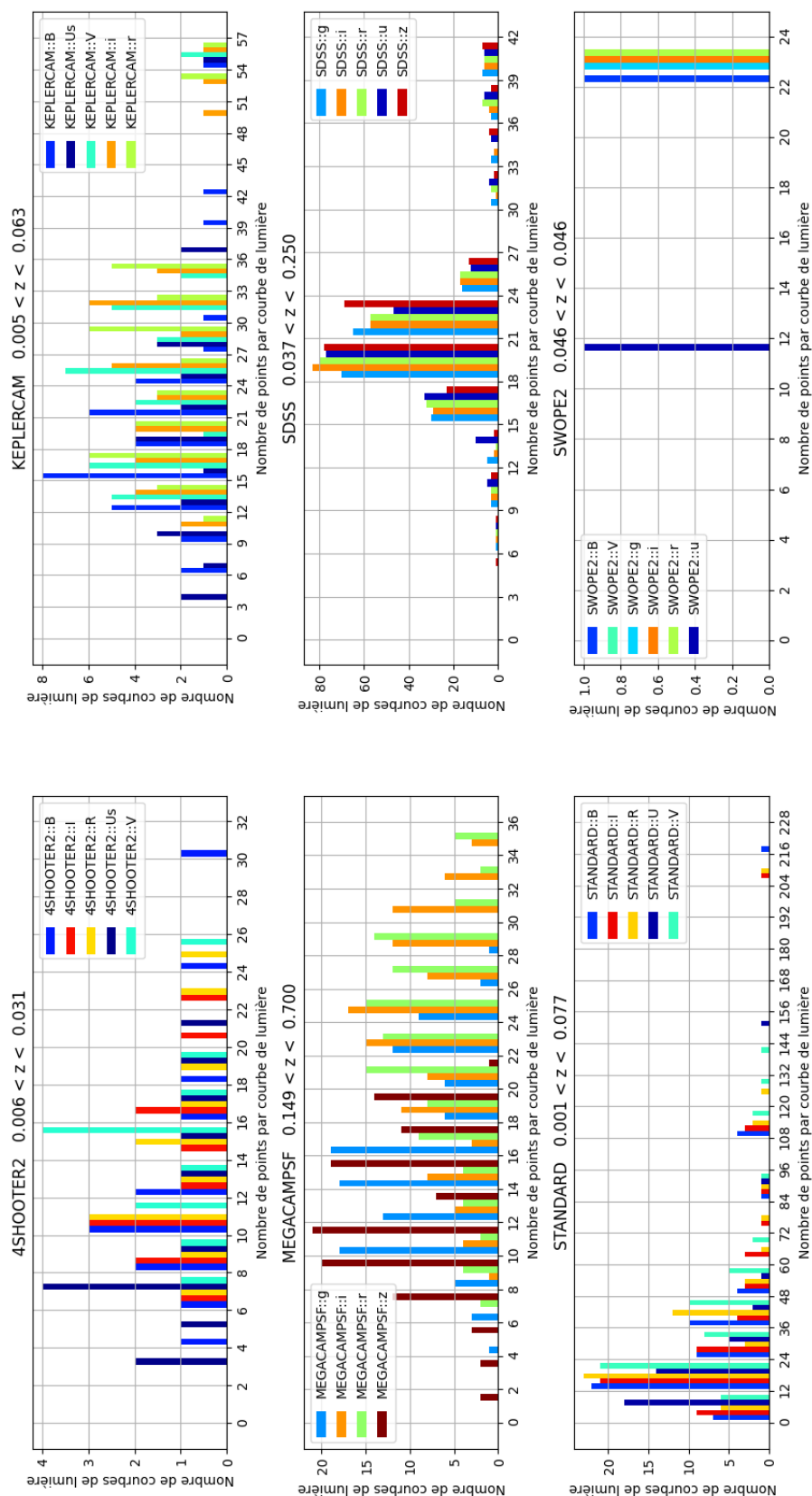


FIGURE 7.8 – Nombre de points par courbe de lumière en fonction des relevés composant JLA.

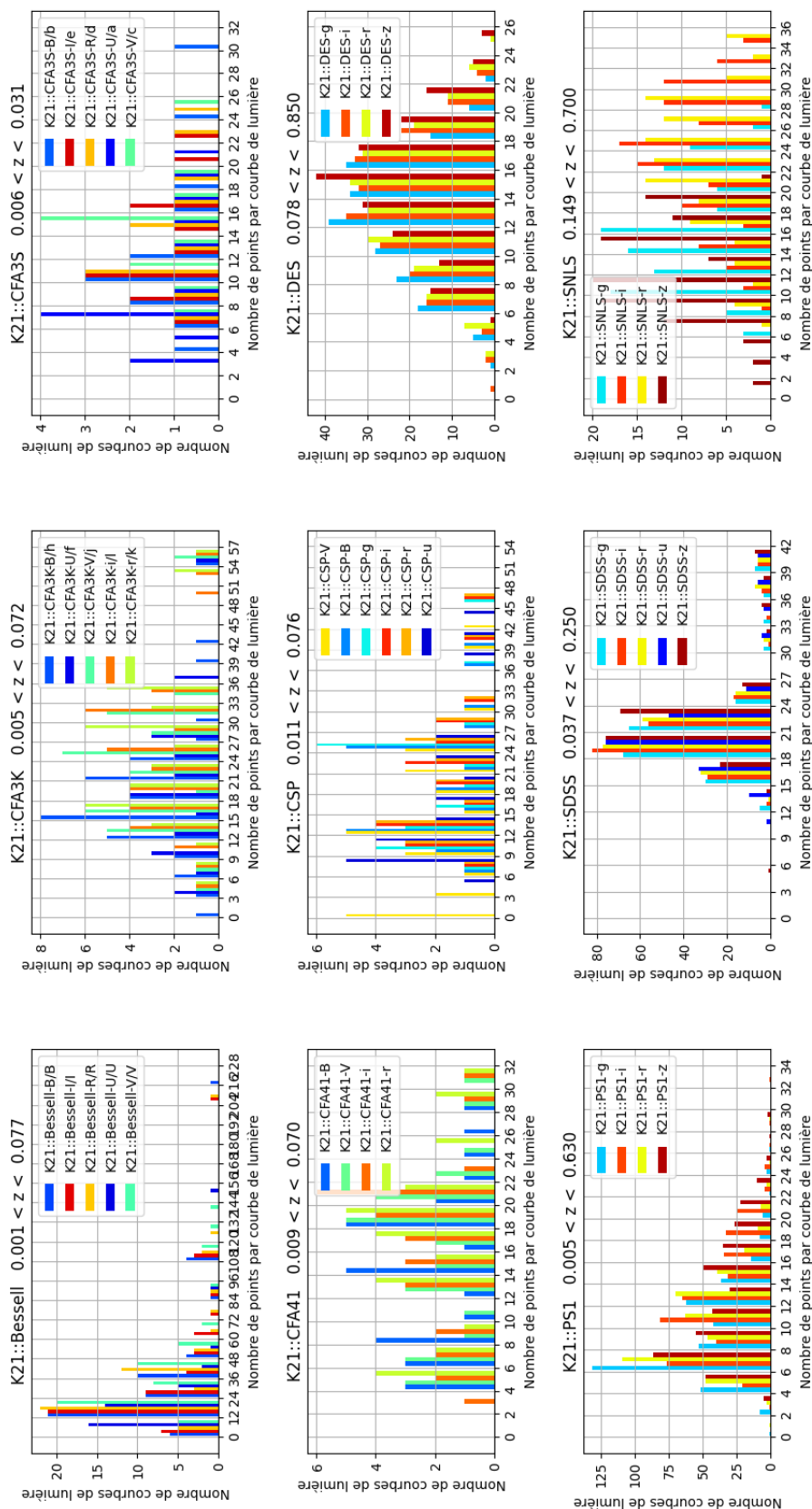
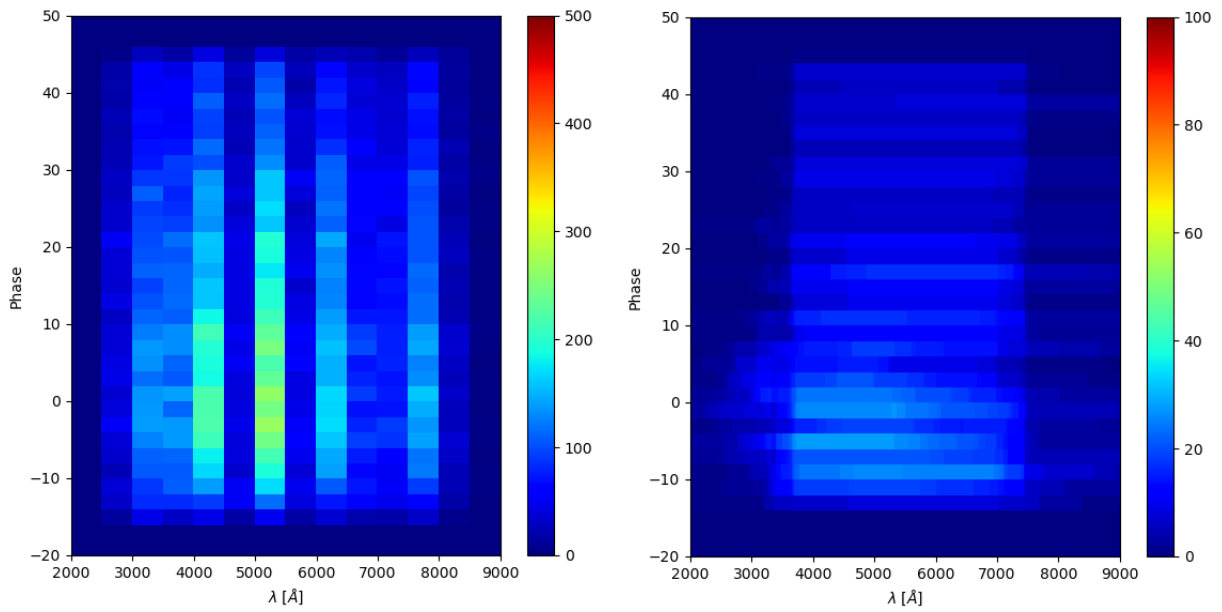


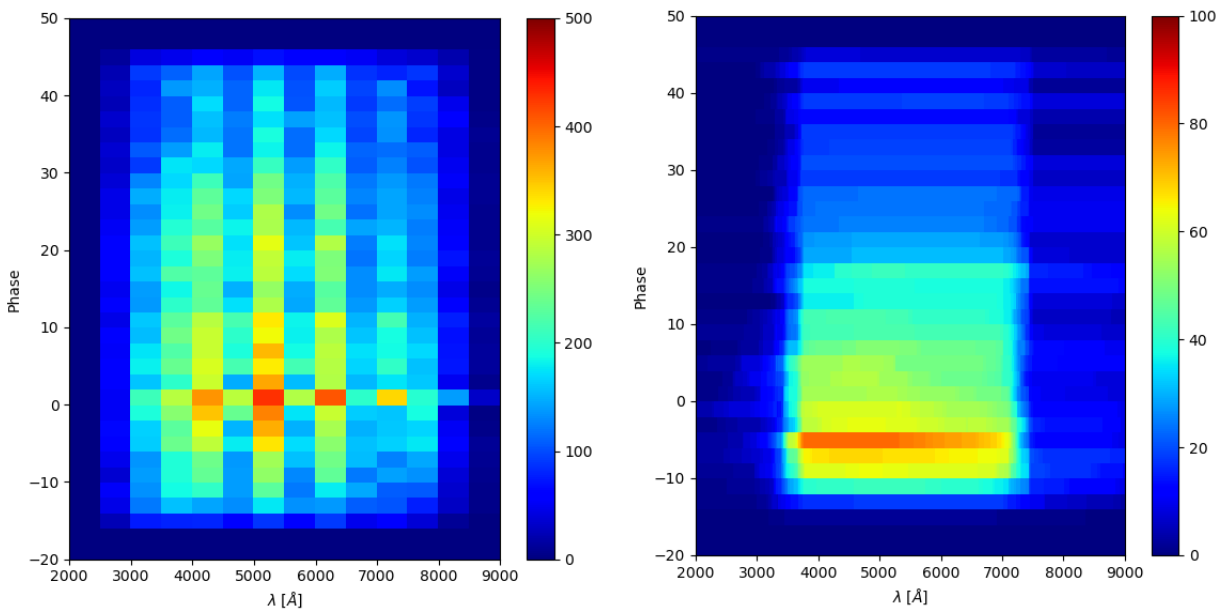
FIGURE 7.9 – Nombre de points par courbe de lumière en fonction des relevés composant K21.



(a) Diagrammes de couverture pour les courbes de lumière du lot d'entraînement de JLA, chaque point de photométrie est compté une fois dans la parcelle correspondante.

(b) Diagramme de couverture pour les spectres du lot d'entraînement de JLA, chaque spectre, présentant des points dans une même parcelle est compté une seul fois.

FIGURE 7.10 – Diagramme de couverture des données photométriques et spectroscopiques de JLA.



(a) Diagramme de couverture pour les courbes de lumière du lot d'entraînement de K21.

(b) Diagramme de couverture pour les spectres du lot d'entraînement de K21.

FIGURE 7.11 – Diagramme de couverture des données photométriques et spectroscopiques de K12.

7.2 Flux

Nous passons maintenant à la description de la seconde étape de notre pipeline de simulation. Une fois les cadences simulées ou extraites des lots JLA et K21 nous simulons les courbes de lumière et les spectres correspondants. C'est la deuxième étape de la figure 7.1. Nous évaluons ces valeurs avec le modèle SALT2.4, c'est à dire avec M_0 , M_1 et CL fixés aux valeurs du modèle entraîné pour JLA. Pour les courbes de lumière observées dans les différents instruments, nous utilisons les modèles de transmission correspondants. Cependant certaines observations peuvent sortir du domaine de définition du modèle, par exemple, si la SN est à trop grand *redshift*. On sélectionne donc les données générées afin d'assurer qu'elles peuvent toutes être prédites par le modèle.

Critères de sélection D'abord, on s'assure pour les courbes de lumière et les spectres que le flux est bien défini par l'intervalle de définition des surfaces du modèle. Si ce n'est pas le cas, ces points sont simplement retirés du spectre ou de la courbe de lumière. Ensuite on s'assure que les données appartiennent bien au domaine de définition du modèle. Nous ne voulons pas de phase exprimée dans le référentiel de la SN hors de l'intervalle de phase. Si c'est le cas, pour un spectre ou pour un point d'une courbe de lumière, il est retiré des simulations. La figure 7.12 montre une coupure pour un modèle défini entre 200 et 900 nm. On peut y voir l'intervalle en longueur d'onde de cinq bandes, *u*, *g*, *r*, *i*, et *z* reconstruit dans le référentiel d'émission pour différents *redshifts*. Les lignes verticales rouges représentent les limites de définition du modèle. Si une supernova est simulée à un *redshift* supérieur à 0.6, elle perd sa courbe de lumière simulée dans la bande *u*.

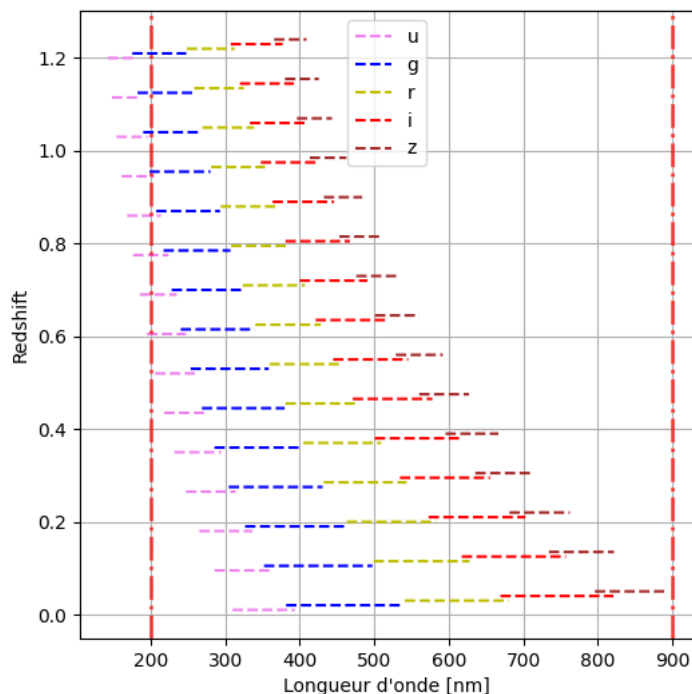


FIGURE 7.12 – Exemple de coupure en longueur d'onde pour un modèle défini entre 200 et 900 nm

7.3 Génération du bruit ajouté

Le dernier ingrédient de la simulation est la modélisation des incertitudes, et la génération du bruit de mesure à ajouter aux valeurs simulées. Le bruit est supposé gaussien, centré et de variance égale à :

$$\sigma^2 = \sigma_{Err}^2 + \sigma_{mod}^2 + \sigma_{seuil}^2 \quad (7.4)$$

Nous allons détailler les différentes composantes de ces incertitudes.

σ_{Err} : Les incertitudes de mesure dépendent du type de données et de la simulation souhaitée.

Pour les simulations de JLA et K21, pour chaque donnée, photométrique ou spectroscopique, le rapport σ_{Err}/ϕ est mesuré et multiplié par la valeur du flux simulé.

Pour les spectres des deux types de simulations idéales, on choisit une incertitude ne dépendant pas du *redshift*, fixée à 2% de la valeur du flux : $\sigma_{Err} = 0.02\phi_{spec}$.

Pour les courbes de lumières, nous considérons deux cas en fonction du type de simulation. Pour les simulations idéales, nous ne considérons que le bruit que de la valeur du flux : $\sigma_{Err} = 0.02\phi_{phot}$. Pour le suivi de LSST, nous avons utilisé un modèle de l'instrument LSST, développé au sein du groupe supernovæ du LPHNE, pour simuler le bruit affectant les courbes de lumières. Le modèle permet d'évaluer la variance du bruit de Poisson affectant les flux des SNe, pour un temps de pose et des conditions d'observation données (*seeing*, fond de ciel, etc). Si on suppose que les flux sont extraits par photométrie de PSF, la variance de l'estimateur de flux σ_F est donnée par :

$$\sigma_{Err}^2 = \left(\sum_i \frac{\psi_i^2}{f\psi_i + f_{sky|i} + \sigma_{RON}^2} \right)^{-1} \quad (7.5)$$

où la somme est effectuée sur tous les pixels voisins de la source. ψ_i est la valeur du modèle de PSF pour le pixel i , f est le flux reconstruit, $f_{sky|i}$ est la contribution du fond de ciel, et σ_{RON} est le rms du bruit de lecture.

L'équation ci-dessus montre qu'il existe au moins deux régimes de fonctionnement : dans le premier, la variance est dominée par les fluctuations de Poisson du flux intégré de la source elle-même. Dans le second, la variance est dominée par les fluctuations de Poisson du fond de ciel. C'est le régime dans lequel nous sommes lorsque nous faisons de la photométrie des SNe.

On a choisi pour notre relevé les deux composantes du relevé LSST, WFD et DDF, (cf. section 3.3.5), différant l'une de l'autre par la valeur du temps d'exposition pour chaque mesure :

- 30 secondes si $z < 0.4$ pour le WFD.
- 600 secondes si $z \geq 0.4$ pour le DDF.

σ_{Mod} : Le bruit lié au modèle d'erreur, σ_{Mod} , correspond à son évaluation à la phase et à la longueur d'onde correspondantes.

σ_{seuil} On ajoute un seuil à l'erreur sur le flux. Pour toutes les simulations, un seuil de 5 pour mille de la valeur du flux simulé est imposé aux données, qu'elles soient photométriques ou spectroscopiques.

Ces dispersions σ sont ajoutées comme bruit aux mesures du flux.

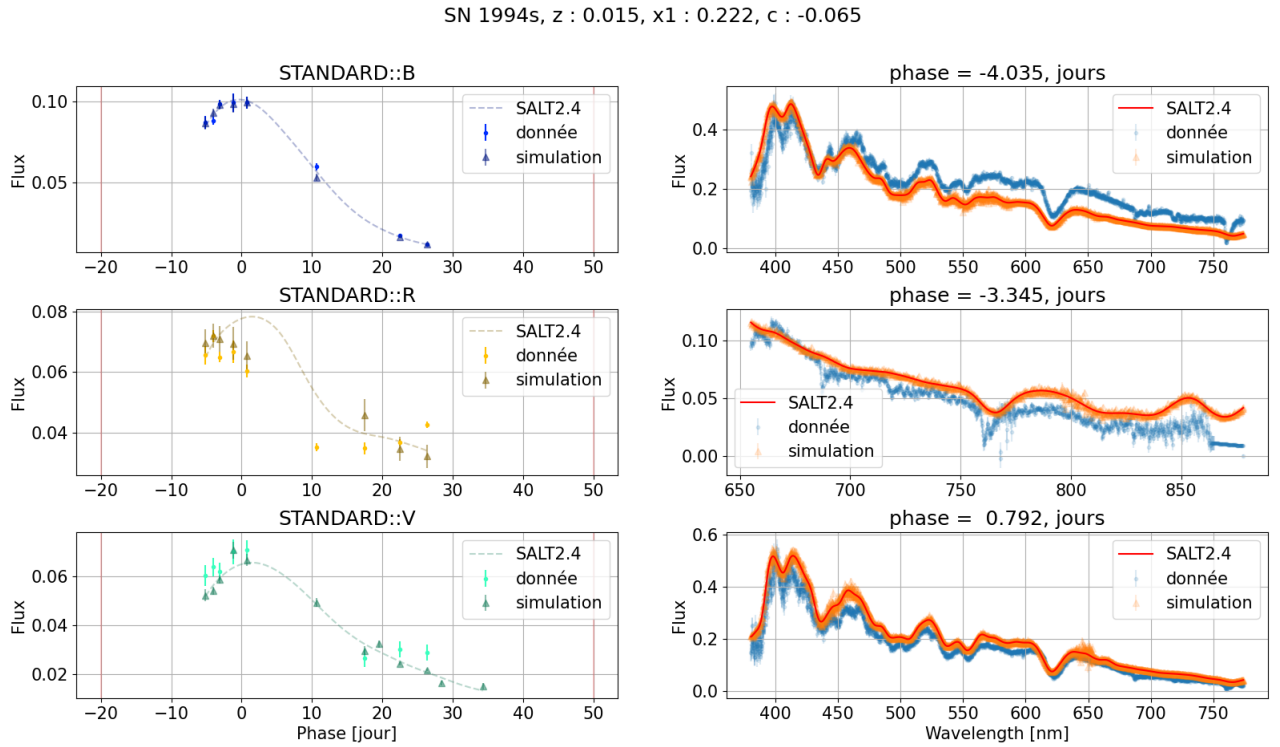


FIGURE 7.13 – Comparaison des données (points) de la SN 1994s à $z = 0.015$, avec les simulations (triangle), pour les données photométriques à gauche et les spectres à droite.

Génération des η : Le terme lié à l'incertitude de calibration est donné par la matrice des incertitudes de calibration \mathbf{V}_η , définie section 6.2.3, dont la diagonale est égale à l'incertitude de calibration de chaque instrument X noté σ_X^2 . Pour chaque bande, on tire un paramètre η_X suivant une loi gaussienne centrée de variance σ_X^2 . Ensuite pour chaque simulation de courbe de lumière nous utilisons le η correspondant. Par défaut, cette valeur est de un pour mille.

Génération de κ : Le dernier terme est lié à la dispersion additionnelle des données autour de la loi de couleur. Nous supposons une distribution de variance $\sigma_\kappa^2(\lambda_r)$, à partir de laquelle on tire pour chaque courbe de lumière un paramètre de dispersion en couleur κ , utilisé pour les simulations.

Chacun de ces éléments est optionnel et peut être activé ou désactivé dans le code de génération, afin de pouvoir prendre en compte ou non le modèle d'erreur, la dispersion en couleur ou la propagation des incertitudes de calibration en fonction des besoins.

7.3.1 Exemple SNe de JLA

Sur les figures 7.13 et 7.14, on peut voir la simulation des données de deux SNe appartenant à JLA. Pour les courbes de lumière, le flux mesuré est représenté par des points, le modèle SALT2.4 estimé en trait plein, et les simulations sont représentées par des triangles, recouvrant le modèle à leur dispersion près. Les points de données présents en dehors du domaine de définition du modèle ne sont pas simulés. Les spectres sont simulés au facteur de re-calibration près.

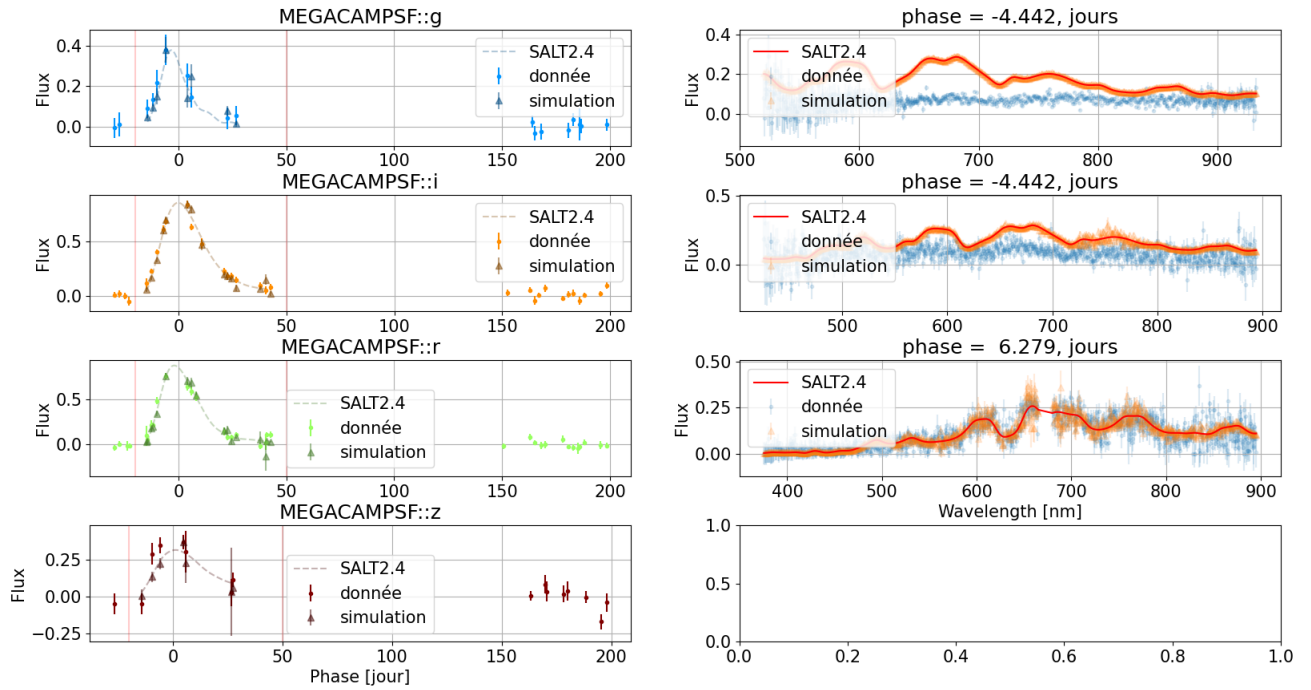
SN 03D1co, $z : 0.679$, $x1 : 0.619$, $c : -0.039$


FIGURE 7.14 – Comparaison des données (points) de la SN05D4av observé par le relevé SNLS à $z = 0.509$, avec les simulations (triangle).

7.4 Conclusion

Nous avons développé un outil permettant de simuler différents lots d'entraînement, du plus naïf avec des erreurs indépendantes du *redshift* et des cadences d'observations choisies, à des simulations plus élaborées simulant le relevé LSST, dont le bruit dépend du *redshift*, jusqu'à des simulations réalistes pour lesquelles la cadence reproduit à l'identique celle des observations réelles.

Les simulations LSST sont utilisées pour tester l'efficacité de la procédure de minimisation, dont les outils ont été développés sur les simulations les plus simples. Les cadences et l'échantillonnage des simulations idéales sont conçues pour imiter autant que possibles celles des ces lots réalistes.

En somme, nous avons créer ces simulations pour pouvoir développer tous les outils dont nous avons besoin. On peut simuler des lots d'entraînement différents pour estimer l'impact d'un entraînement sur des supernovæ différentes, observées à travers des instruments de cadences différentes. Mais il est aussi possible sur une même simulation, de varier les différentes sources de bruit pour comprendre les biais et l'efficacité de l'estimateur. Le chapitre suivant détaille la procédure de minimisation développer pour entraîner le modèle.

Procédure de minimisation

Maintenant que nous avons détaillé, le modèle (chapitre 5), le minimiseur (chapitre 6) et les simulations (chapitre 7), nous décrivons dans ce chapitre l’assemblage de ces différents éléments. Les questions auxquelles nous essayons de répondre ici sont : est-ce qu’un tel ajustement converge ? En combien d’itérations ? Quel est le temps de calcul requis pour atteindre la solution ?

Nous avons construit un minimiseur permettant de contraindre tous les paramètres simultanément. Le modèle est non-linéaire et la dimensionnalité du fit est élevée (quelques dizaines de milliers de paramètres). Le succès d’un tel ajustement dépend en premier lieu de la qualité de l’initialisation du fit. Pour nous assurer de cela, nous déterminons un bon état initial à l’aide de deux minimisations préliminaires. Lors du premier ajustement, nous initialisons le modèle $\{M_0, M_1 \text{ et } CL\}$, les paramètres des courbes de lumière $\{X_0, t_{max}, X_1, c\}$, les paramètres de nuisance η, κ et le prior σ_κ . Dans un second temps, la surface du modèle d’erreur est ajustée à modèle fixé. L’ajustement final est réalisé à partir de cet état initial, tous les paramètres du modèle étant libres.

Dans une première section, nous détaillons la procédure d’ajustement. Ensuite nous décrivons l’ajustement du modèle complet sur les simulations LSST idéales. Enfin, nous présentons quelques résultats pour un lot de données du type JLA.

Contents

8.1	Procédure de minimisation sur les simulations LSST	176
8.1.1	État initial	176
8.1.2	Critères de convergence	177
8.1.3	Détermination des hyper-paramètres	178
8.2	Procédure d’entraînement	179
8.2.1	Premier ajustement d’initialisation	179
8.2.2	Ajustement du modèle d’erreur	183
8.2.3	Fit complet	183
8.3	Temps de calcul	188
8.4	Test sur lot de données réalistes	190
8.4.1	JLA	190
8.4.2	K21	199
8.5	Conclusion	199

8.1 Procédure de minimisation sur les simulations LSST

Nous discutons d'abord un entraînement réalisé sur 500 supernovæ suivies avec LSST selon la procédure décrite à la section 7.1.1.

Génération Lors de la génération, nous fixons l'incertitude sur la calibration des instruments à 0.1%. Les données générées incluent un bruit de calibration : pour chaque réalisation du lot de données, un jeu de paramètres $\boldsymbol{\eta}$ est généré suivant une loi gaussienne centrée en 0 et de dispersion 0.001 et le bruit de calibration résultant $\mathbf{K}\boldsymbol{\eta}$ est propagé aux données.

Le color scatter est défini de façon à avoir le même ordre de grandeur que la loi estimée par SALT2 (cf. Figure 5.6). De même, le pipeline de simulation génère un jeu de paramètres $\boldsymbol{\kappa}$ suivant la loi définie par le prior \mathbf{V}_κ et le propage jusqu'aux données.

Pour définir le modèle d'erreur (*error snake*), on pave le plan (λ, p) avec des splines d'ordre 3, avec 3 noeuds intérieurs, soit $5 \times 5 = 25$ paramètres. Les paramètres de cette surface (γ) (cf. équation 5.42) ont été fixés à 0.05 lors de la génération. Le pipeline de simulation génère un vecteur de bruit à partir de cet *error snake*, et l'ajoute à chaque point de mesure.

On rappelle ici quels sont les paramètres utilisés pour la génération :

- les paramètres du modèle $\{M_0, M_1 \text{ et } CL\}$ correspondent à SALT2.4.
- les paramètres des SNe $\{X_0, t_{max}, X_1, c\}$ sont tirés d'ajustement de SALT2.4 sur les courbes de lumières.
- les polynômes de recalibration spectrale valent 1.
- σ_κ est un polynôme de degré 2 en λ_{red} : $\sigma_\kappa = 0.015\lambda_{red}^2 - 0.02\lambda_{red} + 0.05$.
- paramètres de nuisance : $\boldsymbol{\kappa}$ sont tirés selon une loi normale centrée de variance σ_κ^2 et les $\boldsymbol{\eta}$ suivant une loi normale centrée de variance 10^{-6} .
- les paramètres du modèle d'erreur valent 0.05

Chaque SN a quatre ou cinq courbes de lumière¹ de 20 points et un spectre. Pour 500 SNe, il y a au total 2310 courbes de lumière et 500 spectres, soient 11998 paramètres pour le modèle et 25 paramètres pour le modèle d'erreur.

Reconstruction La procédure de minimisation est la suivante :

1. Initialisation du modèle : ajustement de $\{M_0, M_1, CL, X_0, t_{max}^{B^*}, X_1, c, s, \eta, \kappa, \sigma_\kappa\}$
2. Ajustement du modèle d'erreur $\{\gamma\}$
3. Ajustement simultané de tous les paramètres : $\{M_0, M_1, CL, X_0, t_{max}^{B^*}, X_1, c, s, \eta, \kappa, \gamma, \sigma_\kappa\}$

8.1.1 État initial

Une fois les données simulées, nous définissons un état initial pour le premier ajustement. Pour le moment, nous nous contentons de dégrader les paramètres utilisés pour la génération des données, en les multipliant chacun par un coefficient tiré suivant une loi gaussienne $\mathcal{N}(\mu = 1, \sigma = 0.1)$. Les paramètres η et κ sont gardés nuls et les paramètres de dispersion en couleur sont divisés par 20.

La figure 8.1 montre un exemple de l'état initial du modèle pour une SN. Cet état initial assez fortement dégradé a été choisi pour tester la robustesse de l'ajustement. En pratique pour diminuer le nombre d'itérations, on peut partir d'un état beaucoup plus proche de la réalité. Dans

1. En fonction du *redshift*.

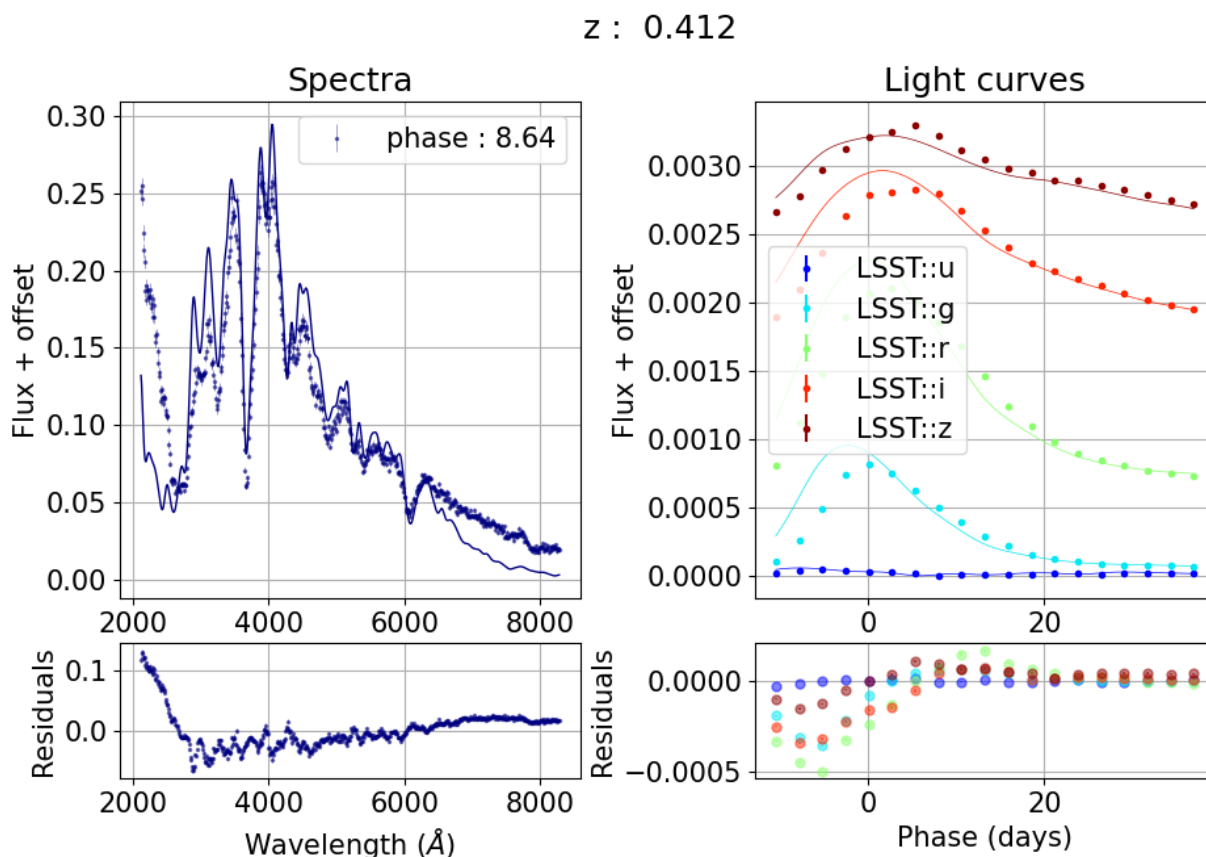


FIGURE 8.1 – État initial du modèle de spectre à gauche (courbe bleu) et à droite pour les courbes de lumière des 5 bandes. Les points représentent les données simulées. La SN est simulée à $z = 0.59$, elle n'a donc que quatre courbes de lumière.

un proche avenir, nous prévoyons de développer une procédure permettant construire un état initial en utilisant un fit rapide des courbes de lumière pour les paramètres $(X_0, X_1, c, t_{max}^{B^*})$. Cette procédure permettra d'éviter d'avoir à conduire les deux ajustements préliminaires décrits plus haut.

8.1.2 Critères de convergence

Nous définissons les critères d'arrêt pour la minimisation suivants : la minimisation s'arrête lorsque

- La différence de $\Delta\chi^2 = \chi_{i-1}^2 - \chi_i^2$ pour une itération i devient inférieure à une limite fixée à 0.1 : $\Delta\chi^2 < 0.1$.
- Ou alors si cette même différence devient négative, auquel cas le minimum du χ^2 correspond aux paramètres estimés à l'étape $i - 1$.

Dans notre cas, la valeur du χ^2 réduit n'est pas un bon indicateur de convergence. En effet, nous ne minimisons pas un χ^2 mais une fonction plus complexe, du fait notamment de la présence des termes $\ln|\mathbf{V}|$. Nous estimons donc la qualité du fit à partir des résidus (i.e. de l'écart du modèle aux données) et des *pulls*, à savoir les résidus pesés par les incertitudes et le modèle d'erreur. Si le fit a correctement convergé, on s'attend à ce que sur nos simulations, les *pulls* soient distribués suivant une loi de Gauss centrée réduite ($\mathcal{N}(\mu = 0, \sigma = 1)$).

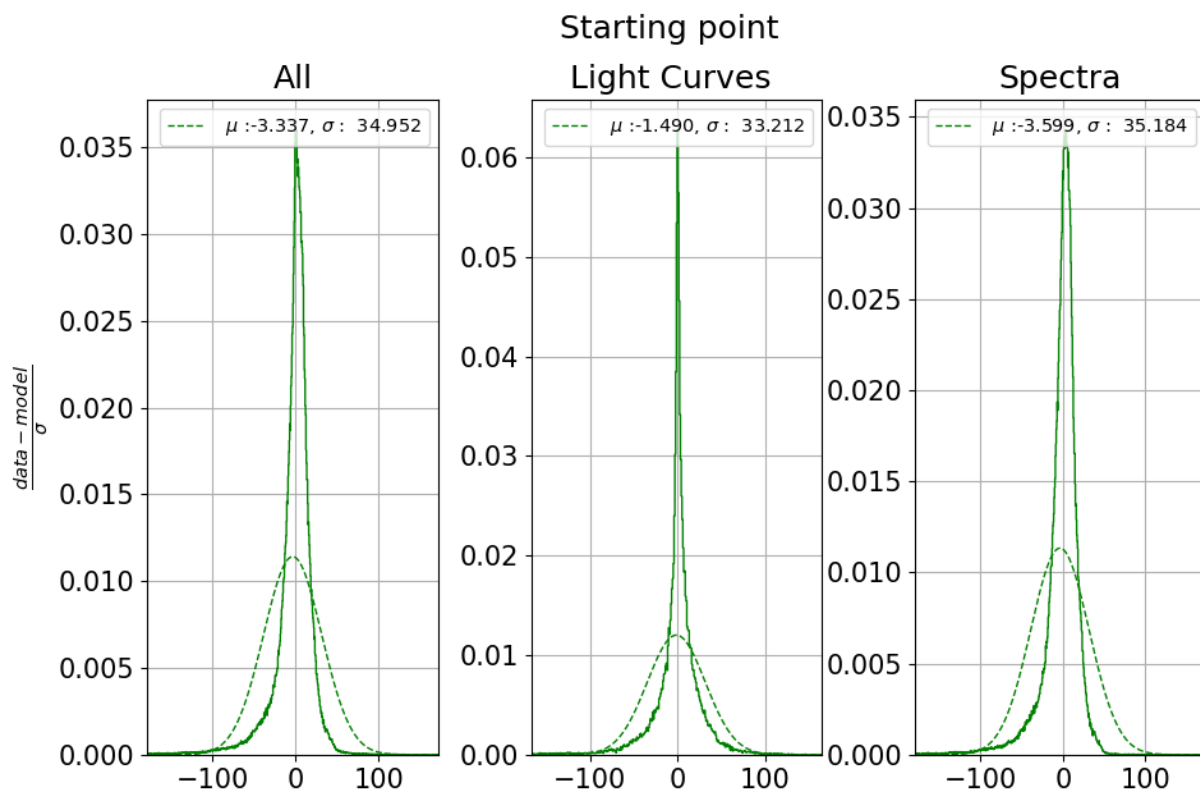


FIGURE 8.2 – Distribution des résidus dé-pesés de l’erreur à l’état initial de l’ajustement. La figure de gauche représente les résidus de l’ensemble des données, celle du milieu des données photométriques et celle de droite des données spectroscopiques. Les courbes représente la distribution d’une loi gaussienne dont les paramètres sont extraits des histogrammes et affichées dans la légende.

$$pull = \frac{y_i - mod_i(\beta)}{\sigma_i} \quad (8.1)$$

où σ_i est la somme quadratique de l’incertitude de mesure et de l’*error snake* si celui-ci est déterminé.

La figure 8.2 présente les *pulls* pour l’état initial. Sur le panneau de gauche sont présentés la distribution des *pulls* pour l’ensemble des données, au centre les *pulls* pour les données photométriques et à gauche pour les données spectroscopiques. Sur cet exemple, nous avons, à l’état initial, des *pulls* très éloignées de la distribution attendue.

8.1.3 Détermination des hyper-paramètres

La valeur des hyper-paramètres, réglant l’intensité de la régularisation et de la pénalité associée aux contraintes, a été déterminé empiriquement. Une limite inférieure est donnée par des critères de convergence : si les pénalités associées sont trop faibles, surtout en début d’ajustement, le fit ne converge pas. Une limite haute peut être déterminée en inspectant les biais éventuels de la reconstruction (selon la méthode décrite au chapitre 9). En pratique, sauf mention explicite, les paramètres de pénalisation quadratique pour les contraintes linéaires ont été fixés à $\mu_{pen} = 10^{10}$. Le paramètre pour la contrainte non-linéaire est fixé à 10^6 . Finalement μ_{reg} est fixé à 100.

8.2 Procédure d’entraînement

8.2.1 Premier ajustement d’initialisation

Comme indiqué plus haut, la procédure d’entraînement débute avec un premier ajustement dont le but est d’obtenir un point de départ raisonnable pour les paramètres du modèle. En pratique, nous avons découvert qu’il était préférable d’inclure dans cette minimisation les variables latentes η et κ . Nous avons également décidé d’y inclure le prior \mathbf{V}_κ . Ce premier ajustement n’est donc pas purement un ajustement du modèle. Une partie du modèle d’erreur est également incluse.

En partant d’une solution initiale assez proche de la solution, pour laquelle les paramètres du modèle sont ceux de SALT2.4, les paramètres des SNe sont dégradés en les multipliant par un coefficient tiré suivant une loi gaussienne $\mathcal{N}(\mu = 1, \sigma = 0.1)$. Ce premier ajustement requiert une dizaine d’itérations. Nous avons vérifié qu’en partant d’un état initial fortement dégradé, tel que celui que nous avons présenté plus haut, l’ajustement converge toujours. Il requiert simplement plus d’itérations – de l’ordre d’une vingtaine pour notre exemple.

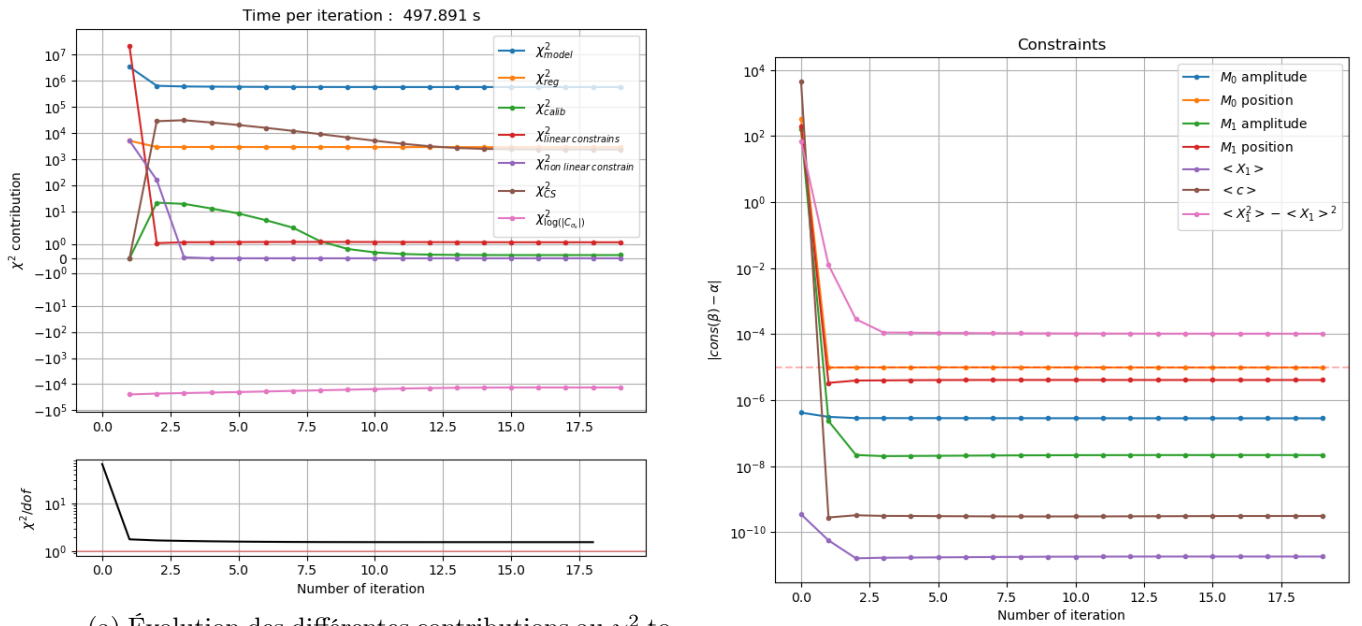
La figure 8.3a montre, sur le panneau supérieur, l’évolution des différentes contributions au χ^2 total, et, sur le panneau inférieur, le χ^2 réduit ($\chi^2/(361684 - 11998)$). Comme attendu, la principale contribution au χ^2 provient du modèle. On peut voir que la contribution de la régularisation n’est pas négligeable, elle est ici inférieure de plus de 2 ordres de grandeurs à la contribution du modèle. La contribution du prior tenant les paramètres de la dispersion en couleur est du même ordre de grandeur. La contribution des contraintes est en revanche négligeable. Ceci est attendu, puisque chaque contrainte brise une dégénérescence exacte du modèle.

Contraintes La figure 8.3b présente l’évolution des différentes contraintes en fonction du nombre d’itérations. Les contraintes sont très rapidement respectées – dès la troisième itération. On vérifie que puisque l’hyper-paramètre μ_{pen} est fixé pour les contraintes linéaires à 10^{10} , chaque contrainte est respectée à 10^{-5} (limite représenté en tirets rouges), et ce dès la première itération. La contrainte sur les variances des X_1 est satisfaite à 10^{-3} près (cf. Figure 8.3b). Le respect de cette contrainte prend plusieurs itérations, probablement à cause de sa non linéarité.

Les figures 8.4a et 8.4b montrent que les contraintes sur les paramètres de couleur et de forme sont parfaitement respectées à l’issue de l’ajustement. La moyenne des paramètres de couleur reconstruits est strictement égale à 0. La moyenne des paramètres de forme X_1 est également nulle, tandis que leur variance est strictement égale à 1.

Ce dernier point a quelques conséquences qui méritent discussion : pour un échantillon donné, on dispose toujours avant entraînement de valeurs X_1 et c pour chaque objet. Si l’on travaille avec des données réelles, ces valeurs peuvent être issues d’un fit des courbes de lumière par un modèle ancien (e.g. SALT2.4). Si l’on dispose de données simulées, ces valeurs sont par exemple issues d’une génération aléatoire. On sera tenté de comparer ces valeurs initiales aux valeurs reconstruites des mêmes paramètres.

En pratique, rien ne garantit que ces valeurs initiales sont centrées. Rien ne garantit non plus que la variance des X_1 sera égale à l’unité. En revanche, à l’issue du fit, par construction, les paramètres estimés respecteront parfaitement ces contraintes, la différence étant intégralement absorbée par les paramètres du modèle. En d’autres termes, la “supernova moyenne” pour laquelle $X_1 = 0$ et $c = 0$ est fonction du lot d’entraînement. Le zéro des couleurs, le zéro et l’échelle des X_1 est une fonction du lot d’entraînement. Nous avons donc la garantie que les couleurs et paramètres de forme reconstruits seront au mieux une fonction affine des paramètres initiaux. Nous avons également la garantie que les paramètres X_1, c d’une même SN appartenant à deux



(a) Évolution des différentes contributions au χ^2 total.

(b) Évolution de la valeur des contraintes.

FIGURE 8.3 – Évolution des différentes contributions du χ^2 et des contraintes pour chacune des 10 itérations.

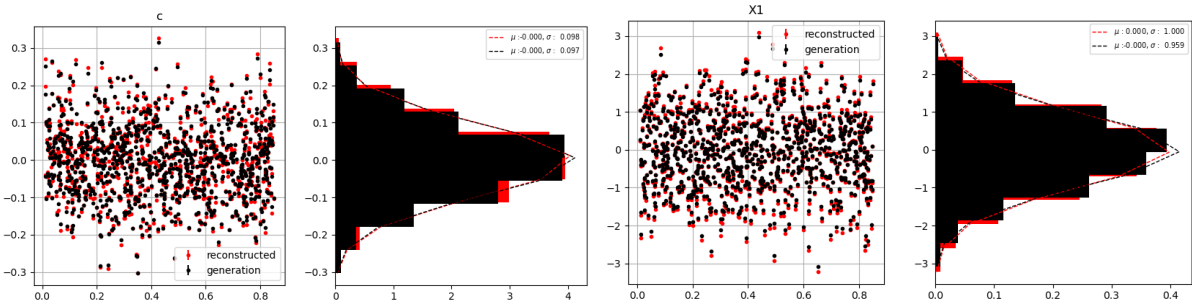
lots d'entraînement différents différeront. Cela n'a pas de conséquence sur la cosmologie, à partir du moment où aucun biais fonction du redshift n'est observable sur les modules de distance standardisés.

Dans la suite, pour simplifier l'étude des biais de reconstruction sur les données simulées, nous nous assurons que les variables X_1 et c générées dans nos simulations sont centrées.

Incertitudes de calibration Les variables latentes η sont également ajustées lors de cette première minimisation. Ces variables n'ont pas d'utilité autre que d'implémenter les degrés de liberté nécessaires à la propagation des incertitudes de calibration. Leur valeur n'a donc pas de signification physique directement interprétable. On montre néanmoins les valeurs reconstruites figure 8.5a. On constate que la variance des η reconstruits est inférieure à la variance spécifiée par le prior associé. Une partie de la variance injectée dans le problème par la calibration est donc absorbée soit par le modèle, soit par les paramètres propres à chaque SN. Cela n'a pas de conséquence notable, à partir du moment où la variance finale propagée dans la matrice de Fisher est correcte.

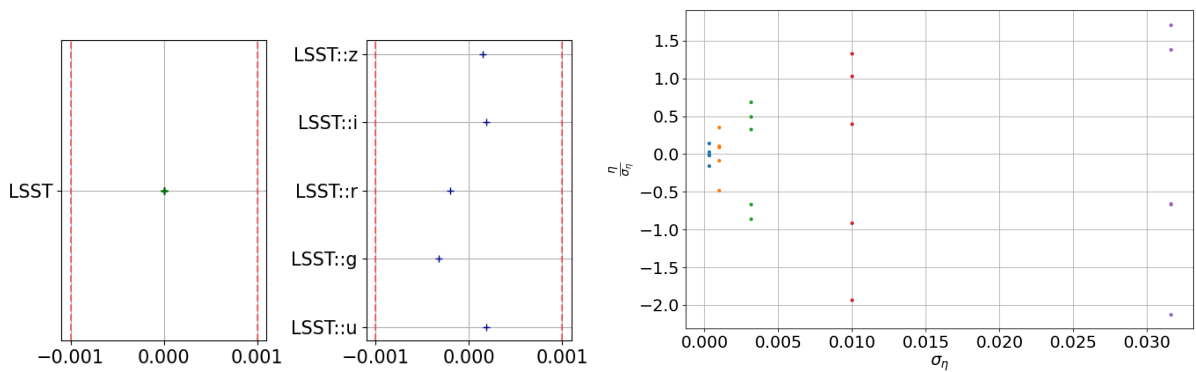
Dispersion en couleur Les variables latentes κ sont également ajustées lors de ce premier fit. Là encore, ces termes ne servent qu'à la propagation de la variance issue du *color scatter*. Les valeurs à l'issue du fit sont présentées sur le panneau central de la figure 8.6, en fonction de la longueur d'onde *restframe*. Les paramètres reconstruits sont figurés en rouge, les valeurs générées en noir. Nous montrons également le prior reconstruit (courbe rouge), et le prior utilisé pour la génération des données (courbe noire). Le panneau inférieur montre les résidus aux paramètres κ générés.

On constate que (1) les paramètres κ sont correctement reconstruits, modulo une plus grande dispersion dans l'UV, due au fait que les incertitudes sur les points de photométrie dans ces



(a) Reconstruction des paramètres de couleur à l'issue de l'ajustement en rouge et à la génération en noir. (b) Reconstruction des paramètres de forme à l'issue de l'ajustement en rouge et à la génération en noir.

FIGURE 8.4 – Reconstruction des paramètres de couleur et X_1 pour 500 simulations LSST.



(a) Reconstruction des paramètres de calibration à la fin de l'ajustement. Pour chaque bande, la valeur η , pour 500 simulations en fonction de la déviation du paramètre est maintenu par la contrainte, dont la valeur à 1σ est tracée en tiret rouge. (b) Reconstruction des paramètres de calibration standard de l'a priori gaussien, la calibration est fixée respectivement à 0.32 ‰, 1 ‰, 3.2 ‰, 1 ‰ et 3.2 ‰

FIGURE 8.5 – Reconstruction des paramètres η pour 500 simulations LSST.

bandes sont plus élevées (2) la forme de la loi de dispersion en couleur est correctement reconstruite (3) la variance issue du fit est néanmoins biaisée (plus faible que la variance générée). Cela est attendu : lorsque l'on ajuste simultanément la moyenne d'un paramètre et sa variance, on s'attend à avoir une variance biaisée, inférieure à la variance vraie. Dans la configuration actuelle, ce biais a un impact sur la reconstruction de l'incertitude affectant les modules de distance, qui sera discuté au chapitre 9.

Reconstruction des paramètres du modèle La figure 8.7 montre les paramètres du modèle à l'issue de l'ajustement (rappel : nous avons recentré les valeurs générées de X_1 et c). On constate que la surface M_0 est correctement reconstruite pour l'ensemble des paramètres contraints le ratio des surfaces est exactement 1. Par contre aux petites et grandes phases, on peut voir des oscillations dépendantes de la longueur d'onde, qu'une régularisation en longueur d'onde pourrait supprimer. La loi de couleur reconstruite est distordue par rapport à la loi de couleur générée. Ceci est une conséquence de l'ajout du *color scatter* et des degrés de libertés introduits par les paramètres η . Cependant, là encore, le seul critère que nous retiendrons pour juger de l'importance ou non de l'effet sera le biais éventuel induit sur les modules de distance. Ceci sera discuté au chapitre 9.

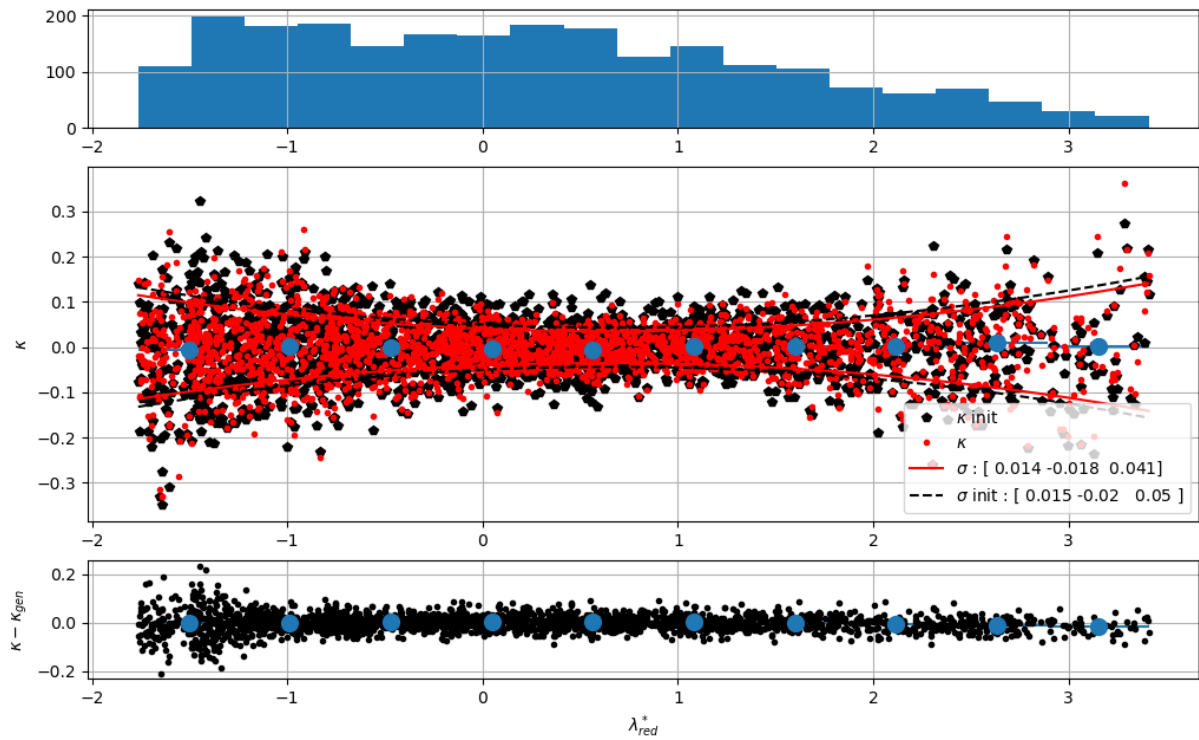
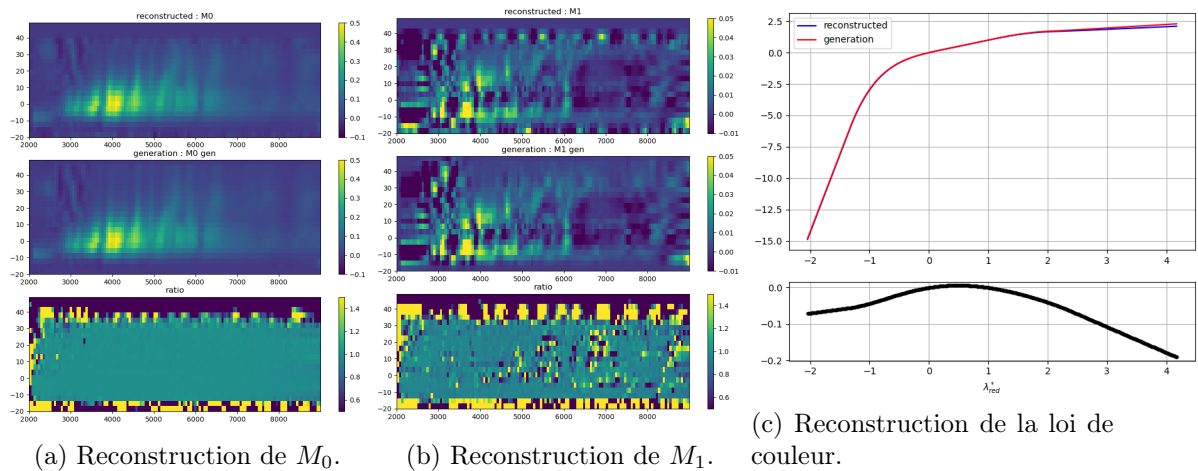


FIGURE 8.6 – Reconstruction de la dispersion en couleur pour une simulation LSST de 500 SNe. Le panneau du haut montre la distribution des λ_{red}^* . Le panneau du milieu représente la reconstruction des paramètres κ et de la dispersion en couleur. Les grandeurs noires sont les valeurs utilisées pour la simulation et les rouges les valeurs reconstruites. Les paramètres de chaque polynôme sont indiqués dans la légende (de gauche à droite, le paramètres d'ordre le plus grand au plus petit). Le panneau inférieur montre les résidus. Les amplitudes des courbes de lumière dans l'UV, $\lambda_{red}^* < -1$ sont bien reconstruites car moins bien mesurées.



(a) Reconstruction de M_0 .

(b) Reconstruction de M_1 .

(c) Reconstruction de la loi de couleur.

FIGURE 8.7 – Reconstruction des paramètres du modèle et comparaison avec les paramètres de SALT2.4. Le panneau du haut montre la surface reconstruite, celui du milieu la surface utilisée pour générer les simulations, celui du bas leur rapport. La loi de couleur reconstruite est affichée en bleu, celle de SALT2.4 en rouge et le ratio des deux est affiché sur le panneau inférieur.

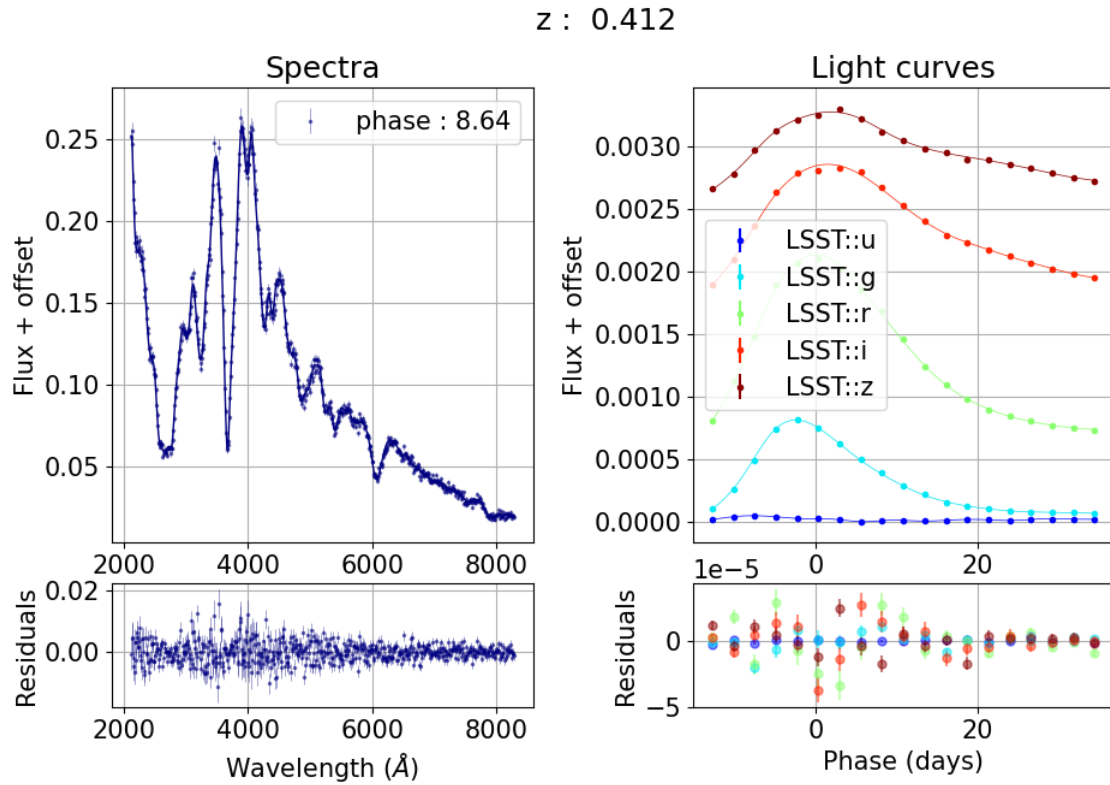


FIGURE 8.8 – Ajustement de notre SN avec le modèle obtenu à l’issue du premier ajustement.

Sur la figure 8.8, on peut voir l’ajustement du modèle sur le spectre et les courbes de lumière de la SN suivie. Les résidus sont maintenant centrés autour de zéro. Néanmoins leur dispersion reste relativement grande par rapport aux erreurs de mesure représentées par les barres d’erreur. La figure 8.9 présente les histogrammes des *pulls* à l’issue de l’ajustement. Ils sont maintenant centrés : le modèle moyen est reconstruit. Par contre, la variance des histogrammes est encore trop grande, signifiant que l’on sous estime la variance des points de données. Pour prendre en compte cette dispersion nous ajustons le modèle de variance, sur les résidus des données et du modèle précédemment ajusté.

8.2.2 Ajustement du modèle d’erreur

Le second ajustement est une initialisation du modèle d’erreur. On ajuste seulement la variance non corrélée des données, paramétrisée par la surface γ . À l’état initial les paramètres du modèle d’erreur sont là encore dégradés, le coefficient tiré selon une loi gaussienne, centré en 1 et de déviation standard égale à 1. Les critères d’arrêt sont les mêmes. On converge dans notre exemple en 2 itérations.

8.2.3 Fit complet

À ce stade, nous disposons d’une solution de qualité suffisante pour pouvoir réaliser l’ajustement complet (tous les paramètres libres). L’ajustement final, prend de l’ordre de sept itérations. La création de la matrice Hessienne est nettement plus complexe (et chronophage) en particulier par le calcul des produits de dérivées du modèle d’erreur par rapport à tous les paramètres (termes de trace dans l’équation 6.56).

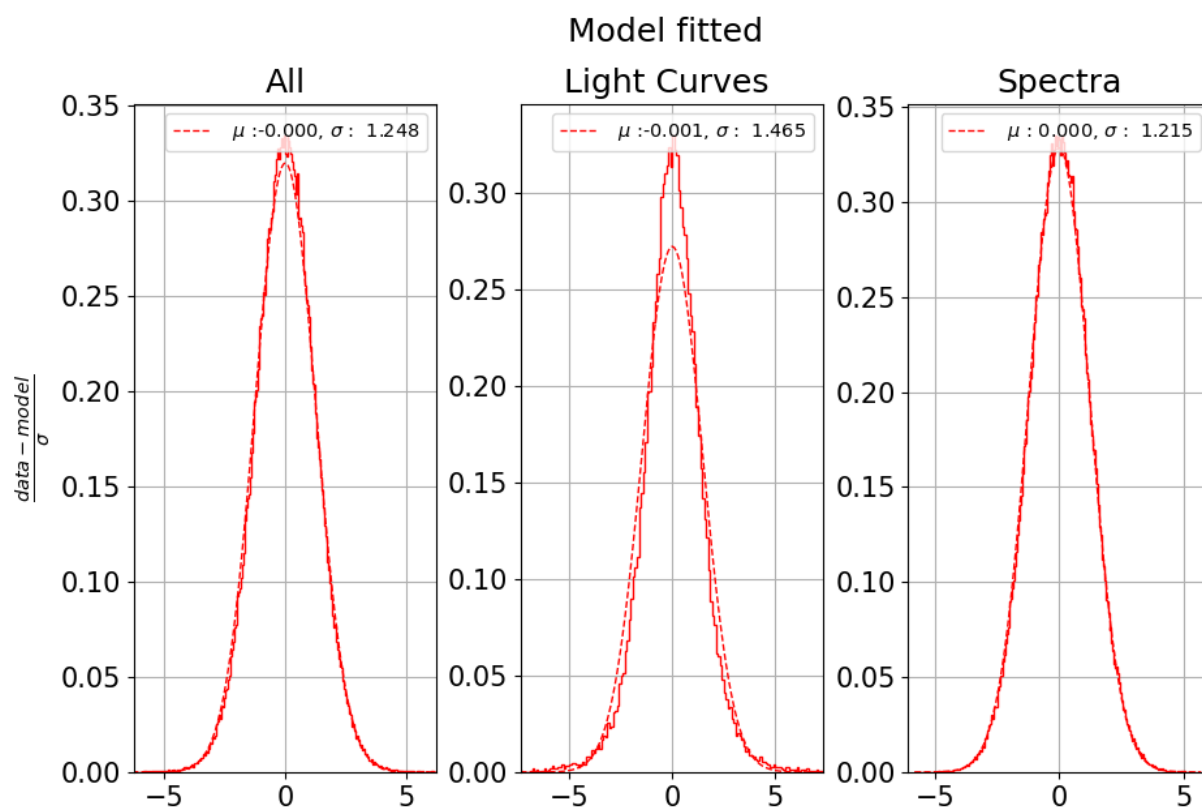


FIGURE 8.9 – Distribution des *pulls* à l’issue du premier ajustement, à gauche de toutes les données, au milieu des données photométriques et enfin à droite des données spectroscopiques. Le trait plein correspond à l’histogramme des *pulls* et la courbe en tiret correspond à une fonction de densité gaussienne dont les paramètres sont indiqués dans la légende. Ils sont centrés mais non réduits, comme attendu puisque seul le modèle et le *color scatter* sont ajustés.

La figure 8.10 montre l’ajustement du modèle, ainsi que le modèle d’erreur à $\pm 1\sigma$ autour du modèle. Sur le graphique des résidus, on peut voir que le modèle d’erreur décrit maintenant la dispersion des données.

Les *pulls* reconstruits avec le modèle d’erreur sont maintenant des gaussiennes centrées réduites (cf. Figure 8.11). Ceci signifie que le modèle moyen est reconstruit et que la variabilité intrinsèque des SNe et ajoutée à la génération a été convenablement reconstruite.

La figure 8.12a montre le ratio de la surface du modèle d’erreur reconstruite par la surface utilisée pour les simulations. Nous voyons que nous pouvons reconstituer la variance non décrite par le modèle seulement aux endroits de l’espace des phases ayant suffisamment de points de données. La figure 8.12b montre l’histogramme des cellules contenant plus de 1% des données, du ratio de la surface du modèle d’erreur et de celle utilisé dans le génération, à l’état initial, en orange, et à l’issue du troisième ajustement, en bleu. La surface est reconstruite à mieux de 95% dans les zone d’intérêt.

La figure 8.13 montre les *pulls* de l’ajustement de toutes les courbes de lumière par bande d’observation et la figure 8.14 montre ceux des spectres par intervalle de phase, ils sont tous centrés et réduits.

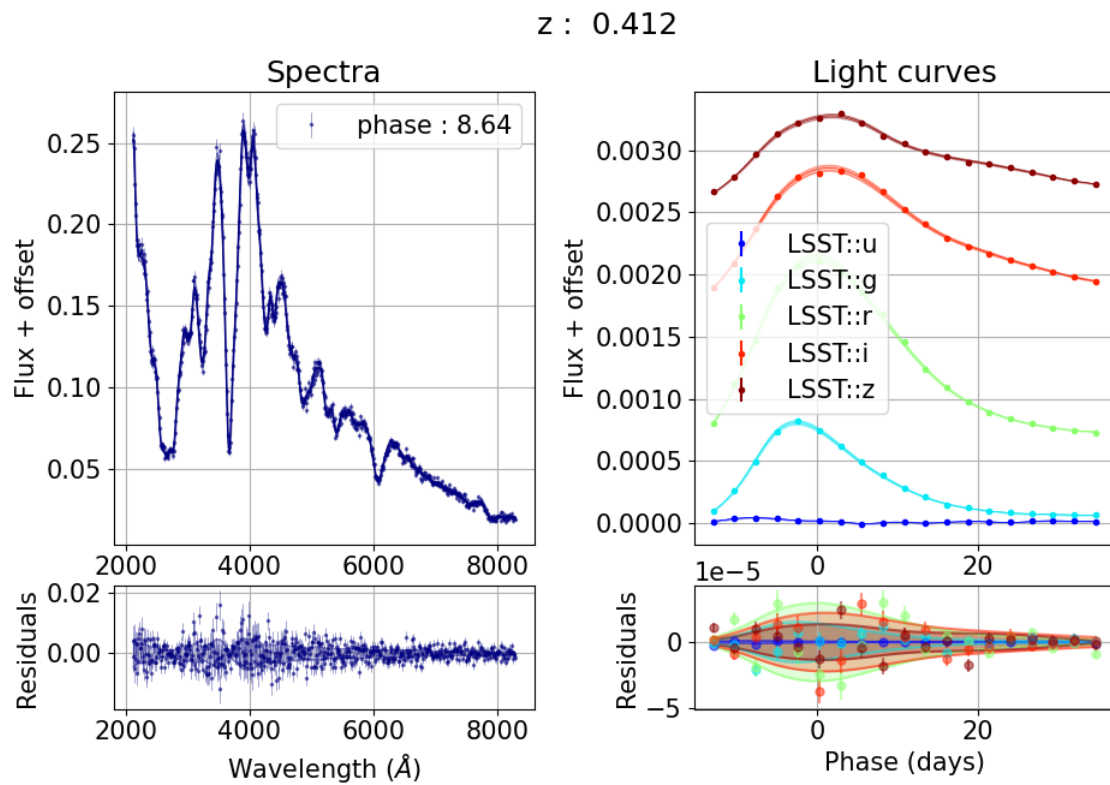


FIGURE 8.10 – Ajustement du modèle et du modèle d’erreur obtenus à l’issue du troisième ajustement sur notre SN.

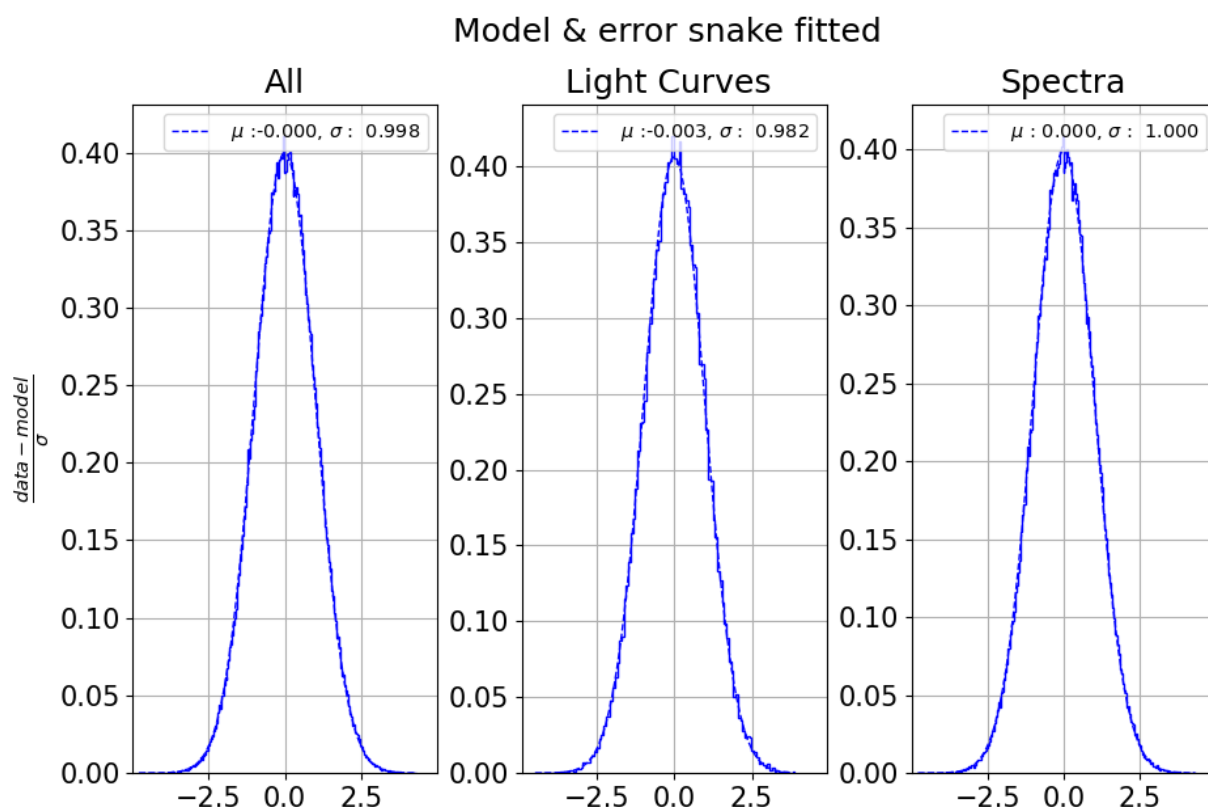
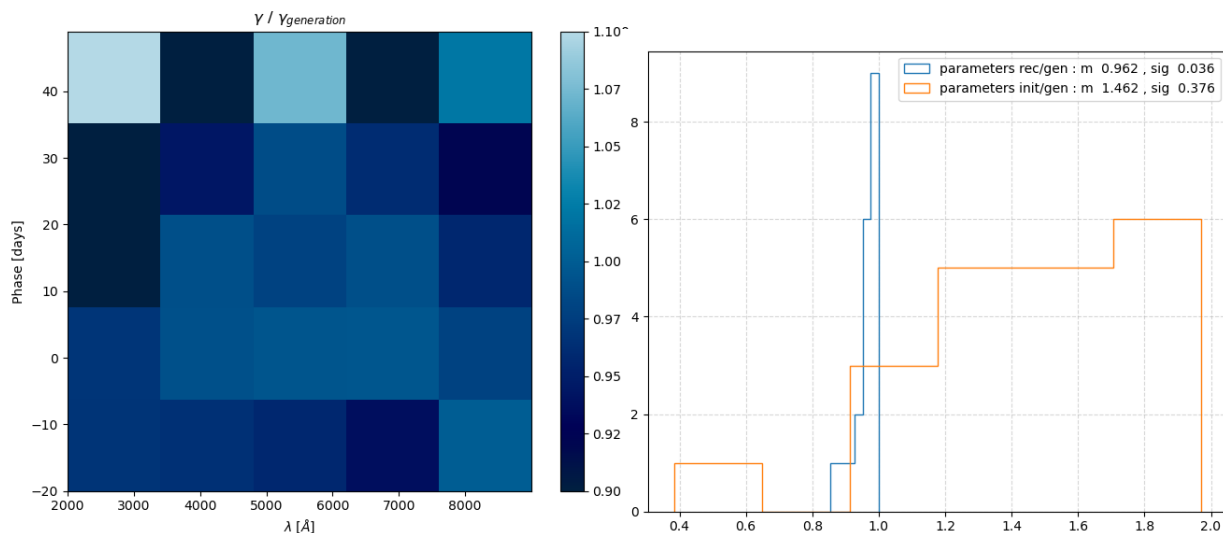


FIGURE 8.11 – Distribution des *pulls* à l'issue du troisième ajustement. Ils sont maintenant centrés et réduits.



(a) Ratio de la surface du modèle d'erreur par la surface utilisée lors de la simulation. Dans les régions, à grande phase et dans l'ultraviolet, les simulations n'ont pas de données. Le modèle d'erreur n'y est pas contraint.

(b) Histogramme des parcelles du ratio de surfaces 8.12a à l'état initial (orange) et final (bleu), contenant plus de 1% des données.

FIGURE 8.12 – Reconstruction du modèle de la surface du modèle d'erreur.

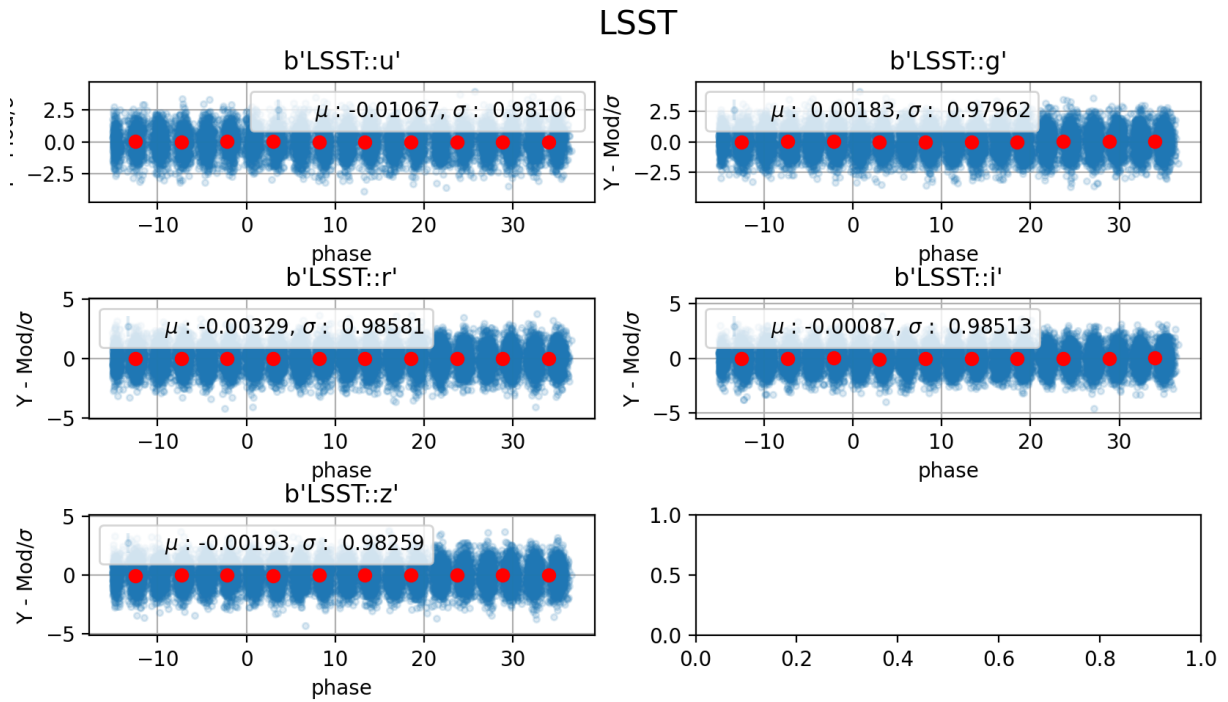


FIGURE 8.13 – *Pull* de l’ajustement du modèle sur les courbes de lumière d’une simulation de 1000 SNe du relevé LSST par bande d’observation.

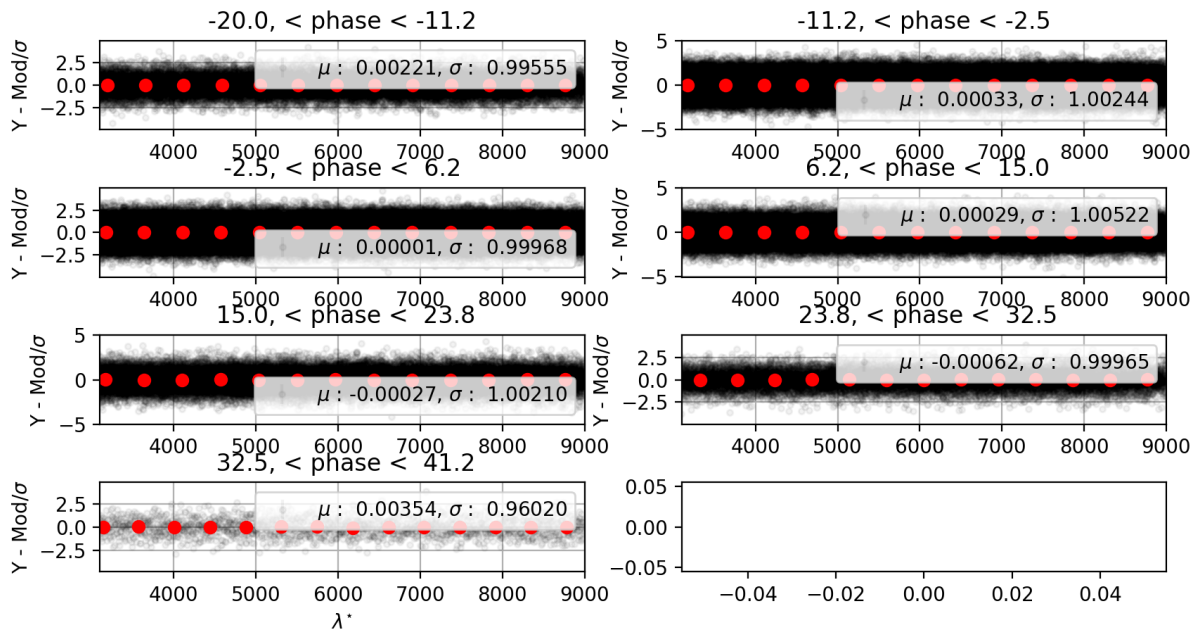


FIGURE 8.14 – *Pull* par intervalles de phases de l’ajustement du modèle sur les spectres d’une simulation de 1000 SNe de LSST.

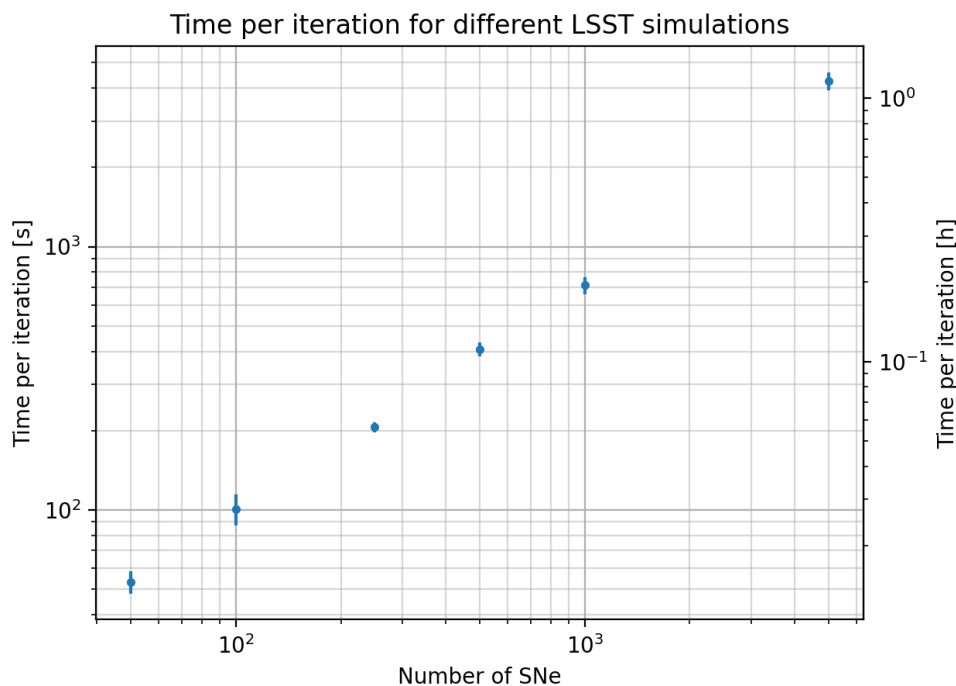


FIGURE 8.15 – Temps de calcul par itération lorsque tous les paramètres sont ajustés simultanément pour 5 lots d’entraînement, respectivement de 50, 100, 250, 500, 1000 et 5000 SNe.

8.3 Temps de calcul

Nous nous intéressons maintenant au temps de calcul nécessaire pour réaliser un entraînement complet. Nous générons à l’aide de notre pipeline de simulation plusieurs lots d’entraînement du relevé LSST de taille croissante (250, 500, 1000, 5000 SNe) ainsi que des lots JLA et K21, et nous étudions le temps passé dans chaque étape de l’entraînement.

Durée d’une itération Notre principale métrique est le temps que prend une itération. Sur la figure 8.15, on présente l’évolution de la durée d’une itération, en fonction de la taille de l’échantillon. On peut constater que jusqu’à 5000 SNe simulées avec un relevé de type LSST, le temps par itération évolue comme le nombre de SNe à la puissance ~ 0.6 . Avec le nombre de SNe, le nombre de paramètres augmente, cependant la densité de la matrice hessienne diminue, passant de $\sim 30\%$ pour 500 SNe, à $\sim 20\%$ pour 1000 SNe et jusqu’à $\sim 8\%$ pour 5000 SNe.

Segments critiques Sur la figure 8.16 nous présentons les principaux segments chronophages observés lors d’une itération, en fonction du nombre de SNe. Les cinq barres correspondent respectivement aux lots d’entraînement de LSST 250 SNe, LSST 500 SNe, JLA, LSST 1000 SNe et K21 et LSST 5000 SNe. Les composantes qui prennent le plus de temps dans le minimiseur sont : l’évaluation du modèle et de ses dérivées, la construction de la matrice hessienne² et la factorisation de la matrice hessienne par l’algorithme de Cholesky. Pour chacune des contributions, les pourcentages par rapport au temps total de l’itération sont indiqués. On peut voir que les lots réalistes prennent moins de temps par itération, simplement car la quantité de données par SN est moindre que dans les simulations LSST-like. La figure montre que pour ces simulations LSST, le principal point critique est la construction de la matrice hessienne, qui correspond à

2. Notamment le produit dans les traces des dérivées du modèle d’erreur par rapport à tous les paramètres.

TABLE 8.1 – Nombre d’itérations, durée d’une itération et temps total passé dans chacun des trois ajustements du processus d’entraînement.

$N_{SNe} - \#_{fit}$	N_{iter}	t_{iter} [h]	t_{total} [h]	N_{iter}	t_{iter} [h]	t_{total} [h]
	Color scatter			No color scatter		
500 - 0	8	0.063	0.504	5	0.061	0.246
500 - 1	2	0.012	0.024	2	0.015	0.031
500 - 2	7	0.113	0.791	4	0.108	0.435
500 - total	17		1.319	11		0.685
1000 - 0	11	0.104	1.142	5	0.098	0.491
1000 - 1	2	0.027	0.054	2	0.033	0.069
1000 - 2	9	0.198	1.59	5	0.145	0.725
1000 - total	22		2.98	10		1.275
5000 - 0	13	0.656	8.530	5	0.500	2.499
5000 - 1	2	0.142	0.284	2	0.135	0.271
5000 - 2	11	1.138	12.529	3	0.890	2.669
5000 - total	26		22.114	10		5.440
JLA - 0	3	0.033	0.099	4	0.040	0.161
JLA - 1	3	0.014	0.042	2	0.017	0.0350
JLA - 2	-	0.054	-	3	0.064	0.191
JLA - total	6		0.141	9		0.416
K21 - 0	3	0.057	0.171	4	0.086	0.344
K21 - 1	4	0.036	0.144	5	0.036	0.180
K21 - 2	-	0.102	-	5	0.116	0.580
K21 - total	7		0.315	11		1.1

plus de la moitié du temps total. Pour le lot JLA, c’est l’inversion de la hessienne. Pour le lot K21, c’est l’évaluation du modèle qui domine.

Temps total d’entraînement La dernière variable d’intérêt est le temps total passé dans les différentes étapes de l’entraînement. Nous présentons ces timings dans la table 8.1, pour différents lots d’entraînement. Comme nous avons constaté que l’ajustement du *color scatter* augmentait significativement le nombre d’itérations, nous présentons nos timings pour des entraînement avec et sans modélisation du *color scatter*. On constate que plus de la moitié du temps est passée dans les ajustements d’initialisation. Obtenir un bon point de départ à moindre coût serait donc extrêmement utile.

Perspectives Notre objectif principal lors de ce développement était de construire un modèle complet, et de montrer qu’il était possible de l’ajuster en une seule minimisation sur un lot de données d’entraînement de bonne qualité. Cet objectif est atteint. L’optimisation du code était considérée comme importante, mais non prioritaire. Au final, les chiffres obtenus sont remarquables. Nous pouvons dès à présent entraîner le modèle sur les lots d’entraînement actuels, JLA en environ une demi heure et K21 en environ une heure. À titre de comparaison, le temps par itération, indiqué dans l’entraînement de SALT3 est de 0.4 heure soit 25 minutes, sur K21. En outre la procédure ne nécessite aucun des entraînements annexes requis pour propager les systématiques de calibration.

Les chiffres présentés dans cette section sont susceptibles d’évoluer assez rapidement alors que

nous finalisons le code d’entraînement. En effet, nous pensons d’abord qu’une procédure d’initialisation optimisée permettra d’éviter les deux ajustements initiaux, ou tout au moins de diminuer le nombre d’itérations requis par le premier ajustement.

Ensuite, le code d’entraînement dans sa version actuelle exploite mal le *multithreading* mis à disposition de manière transparente par la bibliothèque `numpy`. En effet, nous avons décidé initialement de tirer parti du fait que la hessienne du modèle était creuse, en utilisant la bibliothèque `scipy.sparse`. Au final, son taux de remplissage est assez important (30% à 20%) lorsque le nombre de SNe est faible (< 1000). Il diminue ensuite assez rapidement. Or les opérations algébriques sur les matrices creuses sont très mal parallélisables. Par exemple la construction de la hessienne (le calcul de $\mathbf{J}^T \mathbf{W} \mathbf{J}$) est presque exclusivement réalisé sur un seul *thread*. Alors que l’opération équivalente avec la version dense des mêmes matrices est très efficacement multithreadée par `numpy`.

Il se trouve que dans le régime $N_{SNe} \sim O(1000)$, la hessienne reste de taille assez modeste (20 - 40 GB). Elle peut donc tout à fait être stockée sous forme de matrice dense dans la mémoire d’un *worker* des queues à grande mémoire disponibles au centre de calcul. Elle peut aussi être stockée sous cette forme dans la mémoire d’une carte GPU. Les tests initiaux montrent qu’un gain de l’ordre d’un facteur 10 à 50 est possible en réalisant la construction de la hessienne sur une carte GPU. Des environnements tels que `jax`³ permettent de réaliser cela de manière quasi transparente.

Ensuite, la factorisation de Cholesky que nous utilisons n’est pas parallélisée. Des versions multicoeurs de l’algorithme sont disponibles et permettraient là encore de gagner des facteurs importants en temps de calcul.

Enfin, le modèle lui-même peut encore être accéléré : l’évaluation des splines peut être multithreadée sans modification majeure du code. L’utilisation de `jax` permettrait d’accélérer les contractions matricielles au coeur de l’évaluation du modèle et de ses dérivées.

8.4 Test sur lot de données réalistes

Une fois la procédure de minimisation mise au point et testée sur des lots d’entraînement “parfaits”, nous avons voulu la tester sur des échantillons plus réalistes, tels que JLA ou K21. La statistique est plus faible, les courbes de lumière sont moins bien échantillonnées. Parfois même, aucun point de photométrie avant maximum n’est disponible. Enfin les spectres sont affectés d’un bruit de mesure plus important. Nous avons l’assurance que des entraînements itératifs tels que ceux implémentés par SALT2 ou SALT3 pouvaient fonctionner. En revanche, il restait à vérifier que notre ajustement simultané était correctement contraint.

8.4.1 JLA

Sur un échantillon d’entraînement JLA simulé, le premier ajustement converge en 3 itérations et moins de 6 minutes. Le second ajustement (modèle d’erreur) en 3 itérations et de l’ordre de 3 minutes. En revanche, le troisième ajustement (tous les paramètres libres) diverge dès la première itération. Les premières investigations indiquent que les problèmes rencontrés lors du troisième ajustement sont liés à la modélisation du *color scatter*. Il est probable, que le nombre de degrés de liberté introduits par notre modélisation actuelle soit trop important pour pouvoir être contraint par un échantillon tel que JLA. Pour nous assurer que le problème provenait effectivement de la modélisation du *color scatter*, nous avons conduit une série d’entraînements de test, en retirant

3. <https://github.com/google/jax>

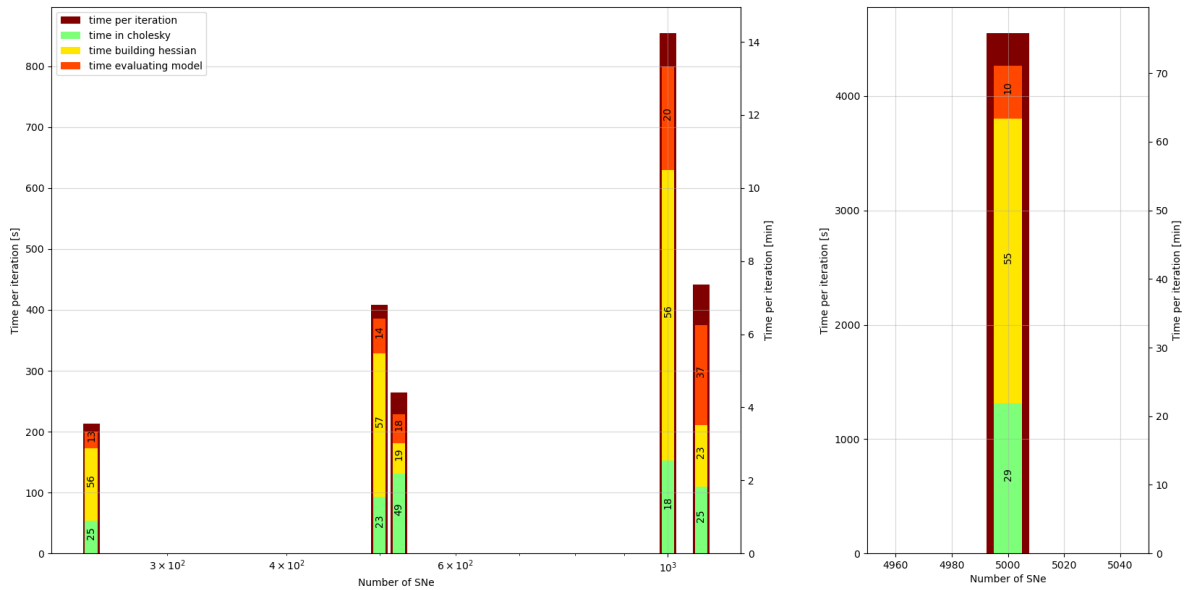


FIGURE 8.16 – Temps de calcul des principales composantes du minimiseur pour différentes simulations. De gauche à droite pour 250, 500, JLA, 1000, K21 et 5000. Les principales composantes sont le temps passé dans l’inversion de la matrice hessienne, dans la factorisation de *cholesky*; le temps passé à sa construction, la taille des matrices rend les produits matriciels lents et le temps passé à évaluer les modèles et leurs dérivées.

la variance additionnelle due au *color scatter* de la génération et de la reconstruction. Comme indiqué table 8.1, tous ces entraînements convergent y compris sur les lots JLA et K21. Pour JLA, l’ajustement final (sans *color scatter*) requiert 3 itérations et 11 minutes au total.

Dans la suite, nous décrivons les résultats des deux premiers ajustements. Nous décrivons le travail en cours liés à la modélisation du *color scatter* dans le paragraphe dédié.

À l’issue de ces deux premiers ajustements, nous constatons que les données sont bien décrites par le modèle. Les figures 8.17 et 8.18 montrent des *pulls* distribués suivant une loi normale centrée-réduite pour les observations photométriques des relevés SNLS et STANDARD par bandes d’observation. La figure 8.19 montre les *pulls* par intervalle de phase de l’ensemble des données spectroscopiques du lot d’entraînement. Pour chaque intervalle, ils suivent une distribution normale centrée et réduite. Les figures 8.22 et 8.23 présentent deux exemples de deux courbes de lumière et d’un spectre ajustés issus de l’entraînement, pour deux SNe, une à bas et une à haut *redshift*.

Timing La troisième barre de la figure 8.16 montre le temps passé par itération dans les différents éléments du minimiseur pendant le troisième ajustement. Une itération complète prend de l’ordre de 4 minutes. L’entraînement complet prenant de l’ordre de 30 minutes. On constate que près de la moitié du temps est passé dans la factorisation de la matrice hessienne. Comme discuté plus haut, ce budget peut être fortement réduit pour cette taille de problème, en repassant à des matrices denses, ce qui permettra d’exploiter efficacement le multithreading et si besoin de porter le problème sur GPU.

Incertitudes de calibration Les incertitudes de calibration affectant les données JLA sont plus élevées 0.3%–0.5% que celles simulées lors de la génération des lots de données parfaits. Par

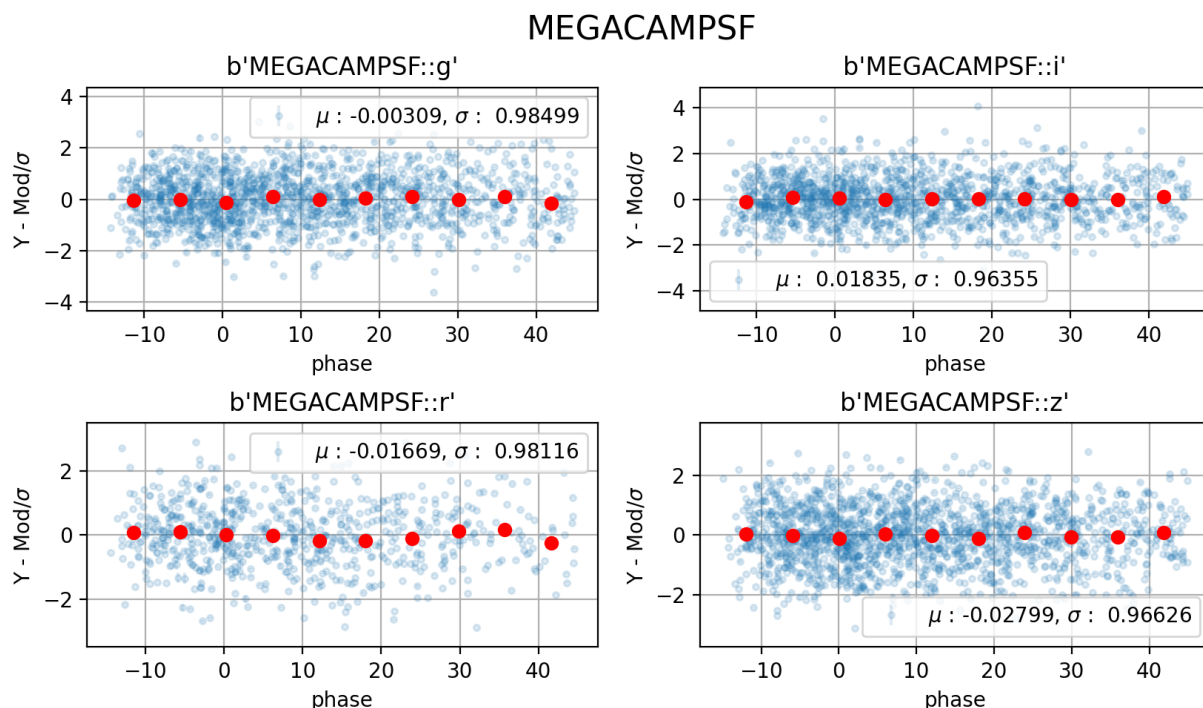


FIGURE 8.17 – *Pull* de l’ajustement du modèle sur les courbe de lumière observées par SNLS de la simulation JLA.

ailleurs, le nombre d’instruments et de bandes passantes en jeu est nettement plus important. Nous avons simulé deux types de lots d’entraînement : un premier, affecté d’un bruit de calibration de 0.1%, et un second avec un bruit de calibration généré suivant la diagonale de la matrice de covariance JLA (cf. table 8.2). Dans les deux cas, nous constatons que l’ajustement converge et que les paramètres η restent contraints (cf. Figure 8.20). Nous constatons là encore que la variance des paramètres η reconstruits est inférieure à ce qui a été injecté dans la simulation. Une partie a donc été absorbée soit dans le modèle, soit dans les paramètres propres aux SNe. Cette fois encore, cela n’a pas de conséquences notables à partir du moment où la totalité de la variance est correctement propagée dans la matrice de Fisher finale. Ceci sera discuté au chapitre 9.

Color scatter On constate en revanche que les paramètres κ introduits pour décrire le *color scatter* ne sont pas correctement reconstruits à l’issue du premier ajustement. (figure 8.21). Nous observons un biais sur la variance des κ reconstruits bien plus important que celui observé sur les échantillons "parfaits". Nous verrons au prochaine chapitre que ce biais affecte directement la variance de fisher des paramètres X_0 .

L’origine de ce biais, et surtout, pourquoi il est aussi important sur le lot JLA, alors qu’il est beaucoup moins prononcé sur les lots “parfaits” de même taille n’est pas encore parfaitement comprise. Deux études sont en cours pour en cerner l’origine. La première consiste à dégrader progressivement la qualité des lots parfaits pour tendre continûment vers un lot de type JLA, en (1) dégradant l’échantillonnage des courbes de lumière ainsi que le nombre de bandes par SN (2) en baissant le niveau de signal sur bruit (3) augmentant le nombre d’instruments en jeu. Une autre étude en cours consiste à limiter le nombre de degrés de liberté associés à la modélisation du *color scatter*. Au lieu de paramétrer les fluctuations autour de la loi de couleur en utilisant un paramètre par bande et par SN, l’idée est d’utiliser une fonction lisse de la longueur d’onde

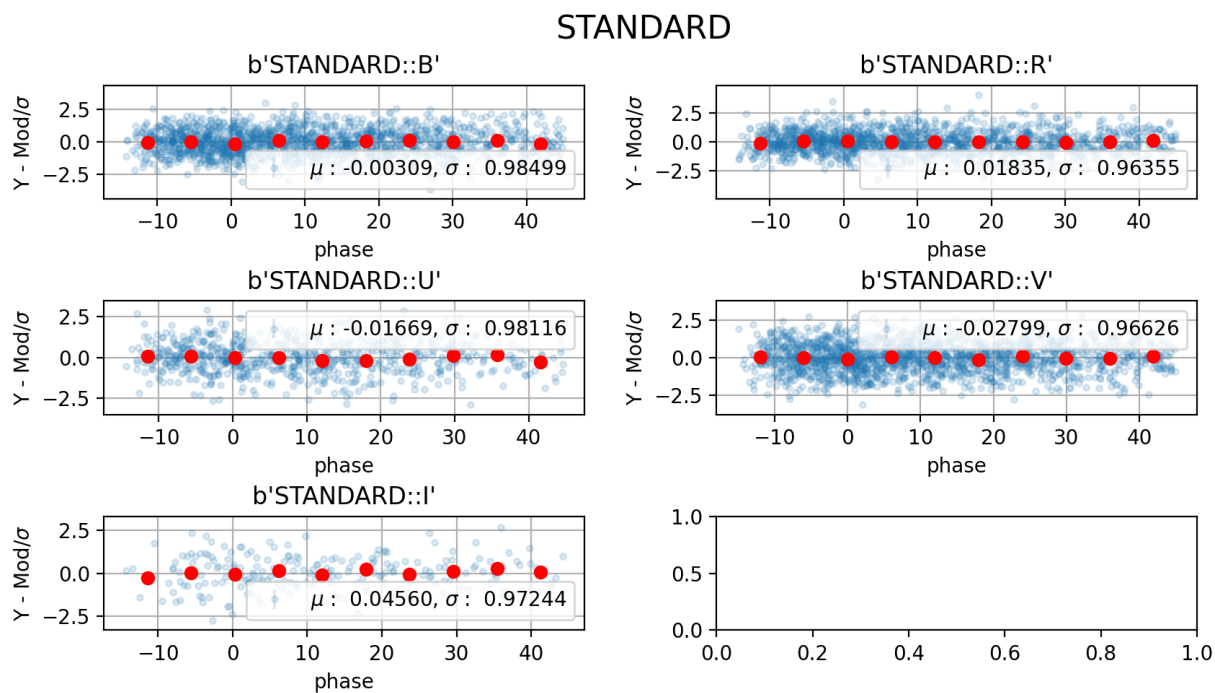
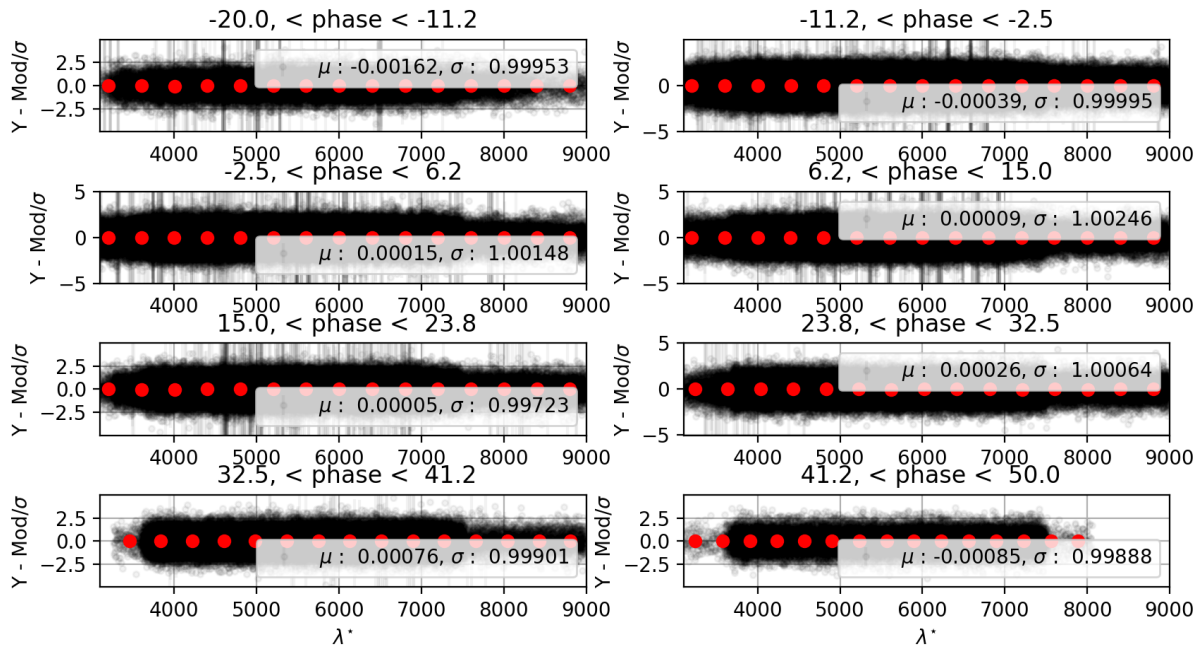


FIGURE 8.18 – *Pull* de l'ajustement du modèle sur les courbe de lumière observées par les relevés STANDARD de la simulation JLA.

restframe de la SN et dépendant d'un nombre plus restreint de paramètres (typiquement 3).


 FIGURE 8.19 – *Pull* par intervalle de phase de l'ajustement du modèle sur les spectres de la simulation JLA.

Instrument	σ_{ZP} [mmag]
STANDARD : :u	100
STANDARD : :g	15
STANDARD : :r	15
STANDARD : :i	15
STANDARD : :B	15
SDSS : :u	8
SDSS : :g	4
SDSS : :r	2
SDSS : :i	3
SDSS : :z	5
4SHOOTER2 : :Us	70
4SHOOTER2 : :B	11
4SHOOTER2 : :V	7
4SHOOTER2 : :R	8
4SHOOTER2 : :I	20

Instrument	σ_{ZP} [mmag]
KEPLERCAM : :Us	31
KEPLERCAM : :B	11
KEPLERCAM : :V	7
KEPLERCAM : :r	25
KEPLERCAM : :i	8
MEGACAMPSF : :g	3
MEGACAMPSF : :r	6
MEGACAMPSF : :i	4
MEGACAMPSF : :z	8
SWOPE2 : :u	23
SWOPE2 : :g	9
SWOPE2 : :r	8
SWOPE2 : :i	8
SWOPE2 : :B	8
SWOPE2 : :V	8

TABLE 8.2 – Incertitude de calibration sur les instruments des différents relevés utilisés dans l'analyse JLA.

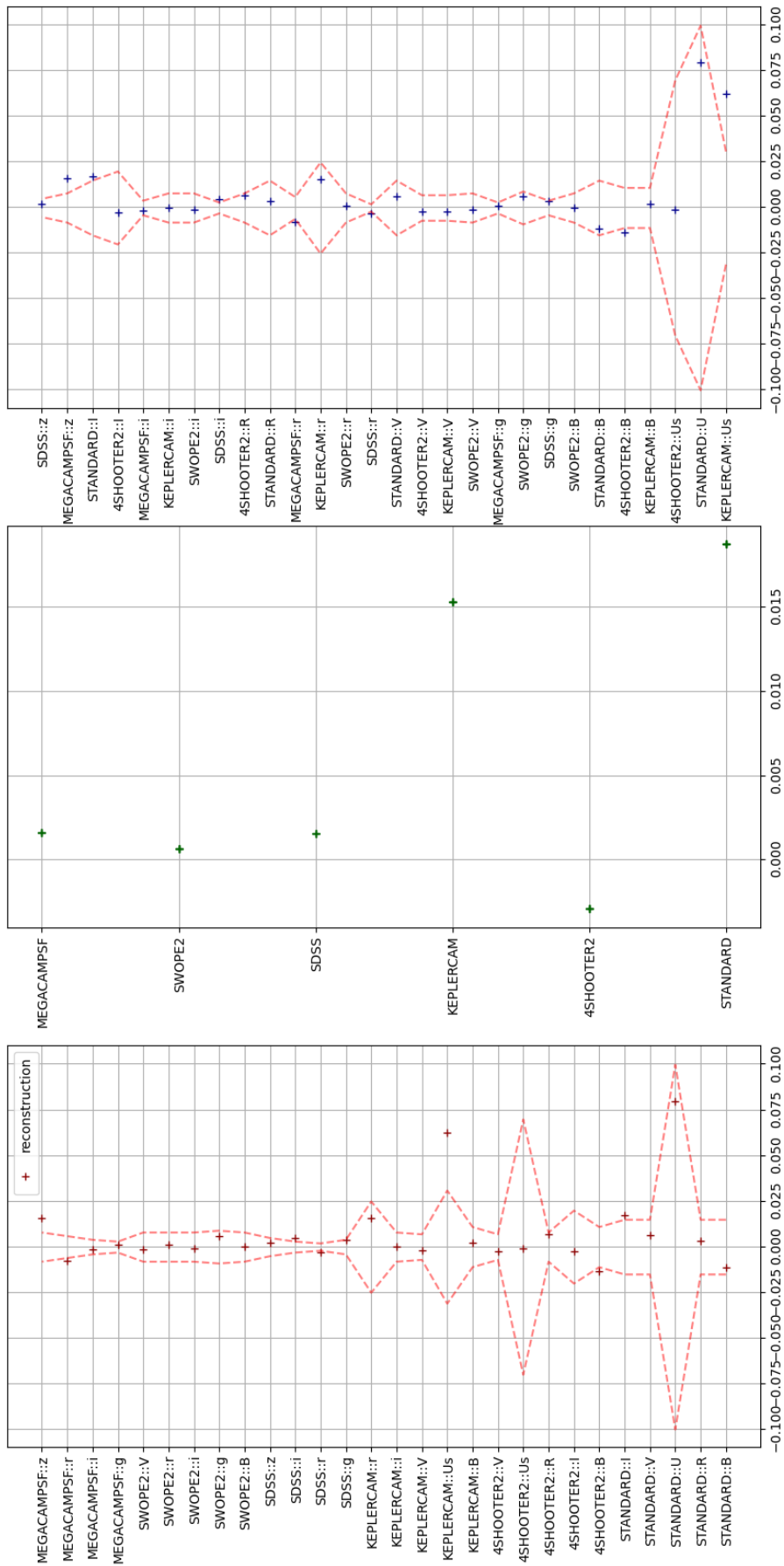


FIGURE 8.20 – Reconstruction des paramètres de calibration η par relevé de JLA sur la panneau de gauche, avec les incertitudes à 1σ en rouge. Sur le panneau du milieu est représenté la somme des paramètres η par relevé. Enfin le panneau de droite montre les paramètres par instrument rangé en couleur.

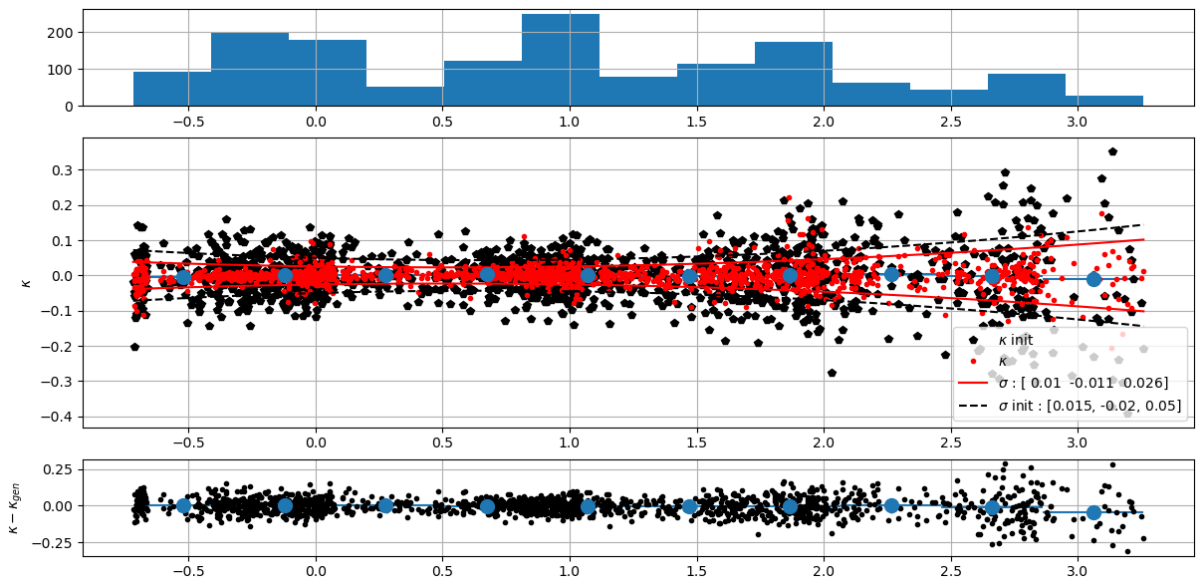
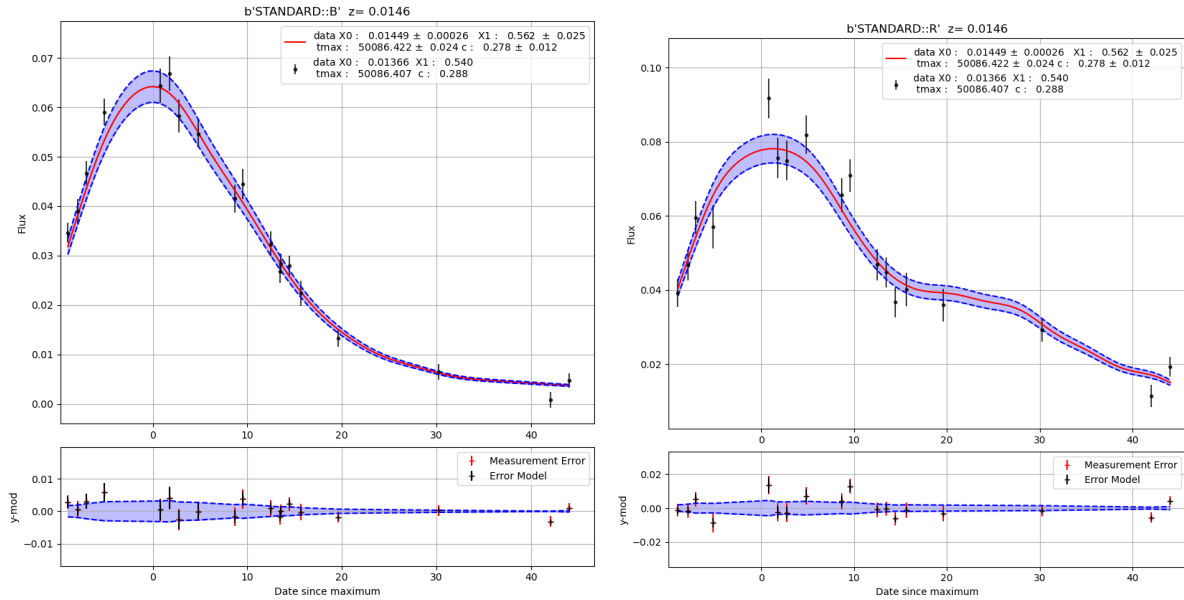
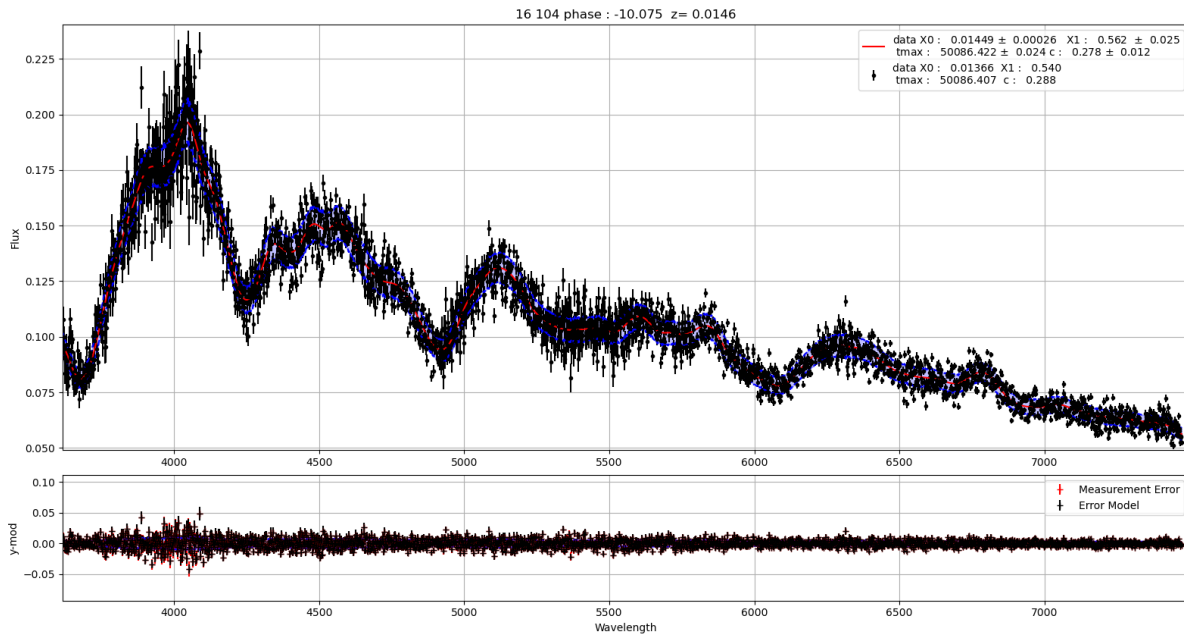


FIGURE 8.21 – Ajustement du color scatter et X_0 sur une simulation JLA, lorsque tous les autres paramètres sont fixés à la vérité.

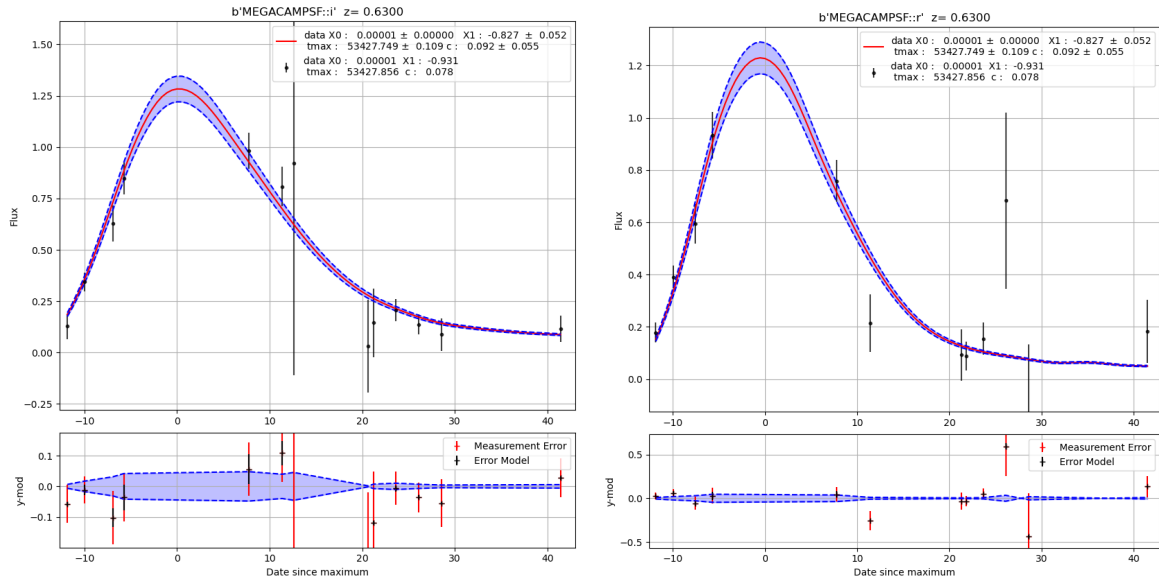


(a) Simulation et reconstruction de la courbe de lumière observés en bande B . (b) Simulation et reconstruction de la courbe de lumière observés en bande R .

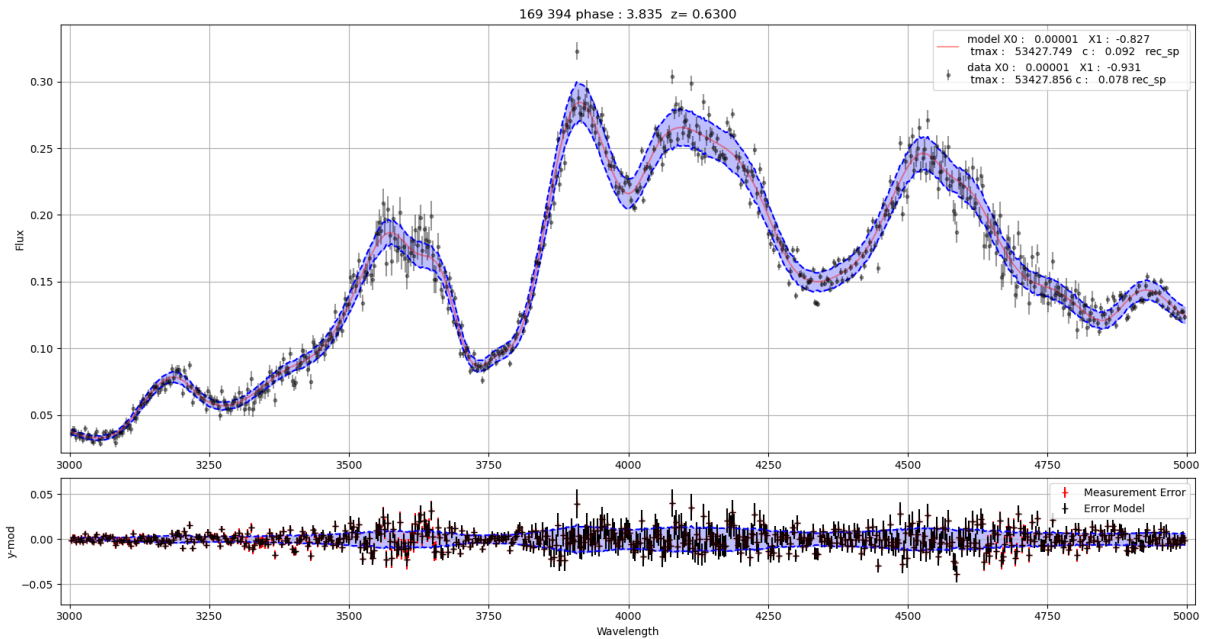


(c) Simulation et reconstruction du spectre à $p = -10$ jours, dont la longueur d'onde est corrigée du *redshift*.

FIGURE 8.22 – Simulation de la SN 1995bd observée à $z \sim 0.015$ par les relevés historiques. Reconstruction du modèle et du modèle d'erreur.



(a) Simulation et reconstruction de la courbe de lumière observée en bande i . (b) Simulation et reconstruction de la courbe de lumière observée en bande r .



(c) Simulation et reconstruction du spectre à $p = 3.8$ jours, dont la longueur d'onde est corrigée du *redshift*.

FIGURE 8.23 – Simulation de la SN 05D2bv observée à $z \sim 0.63$ par SNLS. Reconstruction du modèle et du modèle d'erreur.

8.4.2 K21

La procédure d’entraînement donne sur l’échantillon K21 des résultats extrêmement similaires. Les ajustements d’initialisation requièrent 3 et 4 itérations respectivement, pour un temps par itération de 3.5 et 2.2 minutes respectivement. Le fit principal diverge lui aussi dès la première itération. Là encore, l’origine est notre modélisation du *color scatter*. Sans *color scatter*, la procédure d’entraînement converge en 14 itérations au total et ~ 1 h. Là encore, le modèle ajusté décrit remarquablement bien les données. L’entraînement du modèle sur l’échantillon K21 est plus long que sur JLA d’un facteur 2. Cela est dû à la plus grande quantité de données, et notamment à la plus grande quantité de courbes de lumière. L’évaluation des courbes de lumière est effectivement la partie la plus coûteuse en temps de calcul du modèle. On observe le même biais affectant la variance des paramètres κ reconstruits.

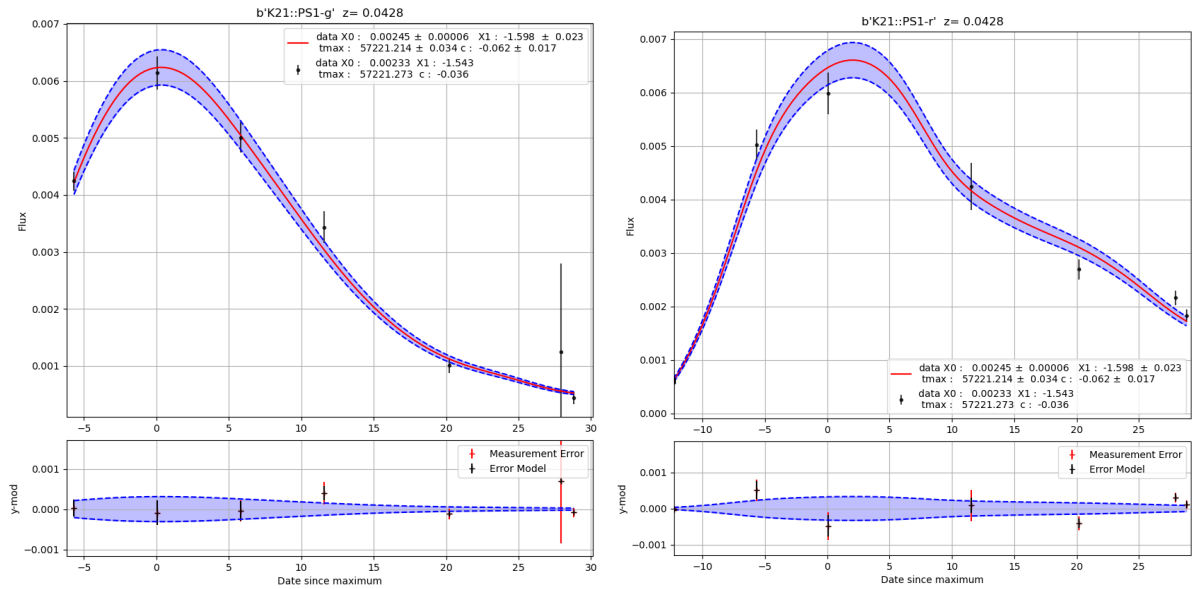
8.5 Conclusion

Nous avons montré dans ce chapitre que le modèle défini au chapitre 5 pouvait être ajusté dans son intégralité, simultanément avec le modèle d’erreur, sur un lot de données de bonne qualité. Il s’agissait de l’objectif principal de cette thèse. Nous avons vérifié que l’ajustement convergait, i.e. que le modèle reconstruit décrivait correctement les données simulées et que les pulls étaient centrés et avaient la distribution attendue.

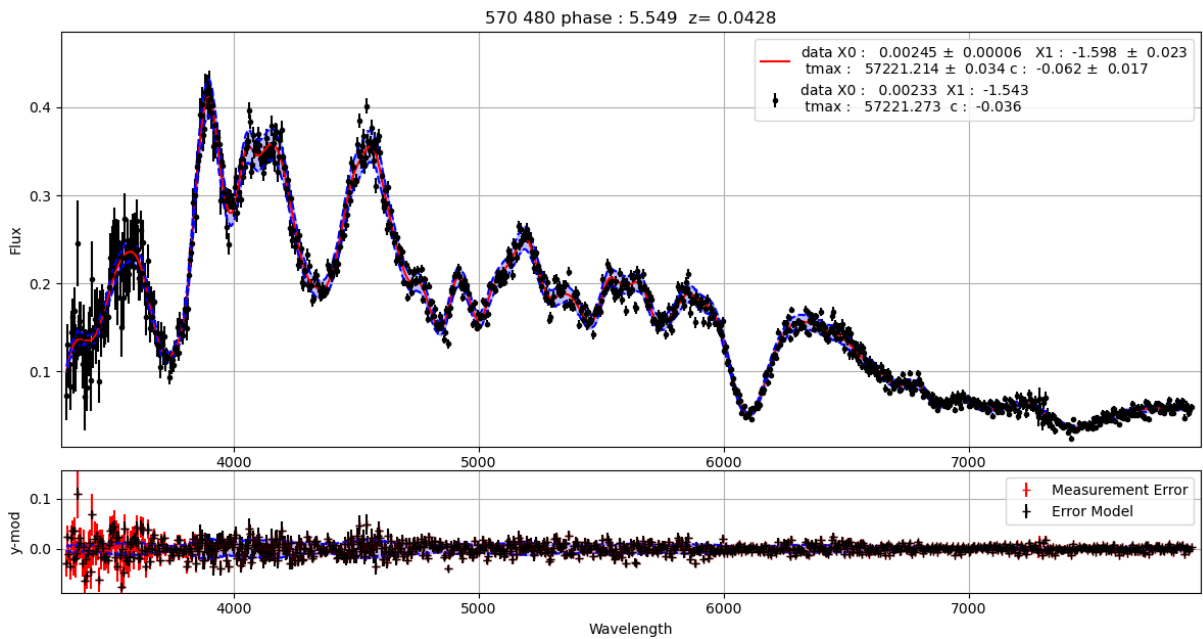
Un secteur du modèle pose encore problème : la modélisation du *color scatter*. Sur des lots de bonne qualité (courbes de lumière très bien échantillonnées, beaucoup de bandes par SN), nous observons un biais affectant la variance reconstruite. Sur des lots simulés reproduisant les lots actuels, notre modélisation du *color scatter* introduit probablement trop de degrés de liberté dans le problème, et fait diverger l’ajustement final. Une piste est explorée pour rigidifier cette partie du modèle.

La procédure d’entraînement repose sur 3 ajustements, les deux premiers ayant pour but d’atteindre un point initial de bonne qualité. L’entraînement est rapide : 30 minutes pour JLA (sans *color scatter*), 1h pour K21 (sans *color scatter*) et de l’ordre de 1 à 3h pour des lots de données LSST-like de taille équivalente mais plus fournis en données. La stratégie d’entraînement peut être optimisée. En particulier il est probablement possible de remplacer les deux ajustements initiaux par un fit SALT2 rapide et simultané des paramètres propres aux SNe. Cela permettra de regagner une fraction substantielle du temps passé dans l’initialisation, quitte à allonger légèrement le fit principal. Par ailleurs, l’examen des segments critiques de notre code a montré que l’essentiel du temps CPU était consacré à deux calculs d’algèbre linéaire : la construction de la matrice hessienne et sa factorisation de Cholesky. Ces deux calculs peuvent être significativement accélérés d’une part en repassant à des matrices denses (ou denses par blocs), ce qui permettrait une meilleure utilisation du multithreading, puis, si besoin, en utilisant les ressources GPU disponibles au CC-IN2P3.

Travailler à l’accélération de l’entraînement se justifie, dans la mesure où nous souhaitons (du moins, dans les analyses de cosmologie) cesser de dissocier la phase d’entraînement du modèle de la phase d’ajustement des courbes de lumière. On y gagne une propagation exacte des incertitudes de mesure, de calibration et de modélisation. À l’issue de l’entraînement, il suffit de conserver les paramètres d’intérêt (X_0 , X_1 , c) et de marginaliser sur tout le reste, modèle compris. Pour juger de la viabilité de cette procédure, nous l’avons implémentée sur des réalisations multiples de nos lots d’entraînement. Cela nous a permis d’une part d’étudier les biais sur les distances standardisées reconstruites, d’autre part, de comparer les incertitudes de Fisher affectant les distances avec leur variance monte-carlo. Cela est l’objet du prochain chapitre.



(a) Simulation et reconstruction de la courbe de lumière observée en bande g . (b) Simulation et reconstruction de la courbe de lumière observée en bande r .



(c) Simulation et reconstruction du spectre à $p = 5.5$ jours, dont la longueur d'onde est corrigée du *redshift*.

FIGURE 8.24 – Simulation de la SN ASASSN-15mg observée à $z \sim 0.0428$ par Pan-Starr. Reconstruction du modèle et du modèle d'erreur.

Études des biais et de l'efficacité de l'estimateur

Dans ce dernier chapitre, nous évaluons la qualité de la reconstruction des paramètres d'intérêt (X_0 , X_1 et c) utilisés pour déterminer les distances standardisées aux SNe Ia. Nous nous appuyons sur le pipeline de simulation décrit au chapitre 7. Pour une cosmologie donnée, nous générons un lot complet (i.e. sans biais de Malmquist associé) de N_{SN} supernovæ de type Ia, dont nous déterminons l'échantillonnage photométrique et spectroscopique. À partir de ces informations, nous générons un nombre arbitraire de réalisations du même relevé. Le but est d'étudier les biais éventuels affectant les distances reconstruites, et de vérifier que les incertitudes de Fisher issues de l'ajustement sont une bonne estimation des incertitudes Monte-Carlo.

Nous décrivons tout d'abord la méthodologie mise en place. Dans la section suivante, nous quantifions les biais affectant la construction des paramètres d'intérêt et des distances. Enfin, nous nous intéressons à la comparaison entre incertitude de Fisher et incertitude Monte-Carlo.

Contents

9.1	Description de la méthode	202
9.1.1	Modules de distance	202
9.1.2	Incertitudes sur les modules de distance	203
9.2	Biais	203
9.3	Variance	210
9.3.1	Tests sur lots de données LSST-like	210
9.3.2	JLA	218
9.4	Conclusion	218

9.1 Description de la méthode

Notre but est double. Tout d'abord, nous voulons comparer les paramètres d'intérêt reconstruits après entraînement $(\hat{X}_0, \hat{X}_1, \hat{c}, \hat{\mu})$ à leurs valeurs "vraies" $(X_{0,\text{true}}, X_{1,\text{true}}, c_{\text{true}}, \mu_{\text{true}})$, générées durant la simulation. Nous voulons également comparer la variance de ces paramètres prédite par l'ajustement, avec la variance empirique calculable sur les réalisations d'un même échantillon.

Notre principal paramètre d'intérêt est bien entendu la distance standardisée reconstruite :

$$\hat{\mu} = -2.5 \log_{10} \hat{X}_0 + \alpha \hat{X}_1 - \beta \hat{c} - M_B \quad (9.1)$$

dont nous voulons comparer (1) la valeur à :

$$\mu_{\text{true}} = -2.5 \log_{10} X_{0,\text{true}} + \alpha X_{1,\text{true}} - \beta c_{\text{true}} - M_B \quad (9.2)$$

et (2) la variance empirique à la variance de $\hat{\mu}$ inférée de la matrice de Fisher de l'ajustement.

Le modèle ajusté est non-linéaire. Nous nous attendons donc à détecter des biais sur X_0 , X_1 et c . Nous nous attendons en revanche à ce que ces biais décroissent au fur et à mesure que la taille des lots d'entraînement augmente. Par ailleurs, comme discuté plus haut, les paramètres X_1 et c sont affectés par des biais "naturels", indépendants de la taille de l'échantillon. En effet, la définition de la SN "moyenne", de couleur et de X_1 nuls est fonction du lot d'entraînement. En comparant les valeurs vraies et reconstruites, nous nous attendons donc à mesurer un biais additif, lié au fait que par exemple, la moyenne des X_1 et c générés n'est pas nulle. Un tel biais n'aura cependant aucun effet sur la mesure cosmologique, à partir du moment où il est indépendant du *redshift*. Ce que nous cherchons à identifier dans la suite, c'est donc l'apparition d'un biais dépendant du *redshift* sur les distances reconstruites.

Pour cette étude, nous nous appuyons sur le pipeline de simulation décrit au chapitre 7. À partir d'une cosmologie d'un lot de SNe générées suivant cette cosmologie, et d'un programme d'observations, nous générons un nombre arbitraire (principalement 50) de réalisations du même lot d'entraînement. Ces réalisations sont identiques. Elles ne diffèrent que par la réalisation du bruit de mesure, du bruit de calibration et de la variance de SN à SN (*error snake* et *color scatter*).

Le modèle est ensuite entraîné sur chaque réalisation du lot d'entraînement, et on compare les valeurs reconstruites et "vraies" des paramètres et des modules de distance standardisés. Plus le nombre de réalisations augmente plus on aura une mesure précise du biais dû à la méthode. On calcule ensuite sur toutes ces réalisations la matrice de covariance empirique des paramètres reconstruits, et on la compare avec l'inverse de la matrice de Fisher (après marginalisation sur les paramètres de nuisance).

9.1.1 Modules de distance

La première étape est de calculer (m_B^*, X_1, c) , puis les modules de distance standardisés à partir de (X_0, X_1, c) . On a construit et normalisé le modèle de façon à ce que :

$$m_{B^*} = -2.5 \log(X_0) \quad (9.3)$$

le calcul des modules de distance standardisés suit :

$$\mu = -2.5 \log(X_0) + \alpha X_1 - \beta c - \tilde{M}_{B^*} \quad (9.4)$$

Dans cette étude, les paramètres de standardisation α et β ne sont pas ajustées mais gardées fixes aux valeurs de JLA, utilisées pour la génération et rappelées ici :

- $\tilde{M}_{B^*} = -19.0906$
- $\alpha = 0.13$
- $\beta = 3$

9.1.2 Incertitudes sur les modules de distance

La matrice de covariance des paramètres $(\hat{X}_0, \hat{X}_1, \hat{c})$ estimés peut être obtenue à partir de l'inverse de la hessienne calculée au minimum. Notons $\mathbf{V}^F = \mathbf{H}^{-1}$, alors la matrice de covariance après marginalisation sur tous les autres paramètres du modèle s'écrit :

$$\mathbf{V}_{\text{SN}} = \begin{pmatrix} \mathbf{V}_{X_0, X_0}^F & \mathbf{V}_{X_0, X_1}^F & \mathbf{V}_{X_0, c}^F \\ \mathbf{V}_{X_1, X_0}^F & \mathbf{V}_{X_1, X_1}^F & \mathbf{V}_{X_1, c}^F \\ \mathbf{V}_{c, X_0}^F & \mathbf{V}_{c, X_1}^F & \mathbf{V}_{c, c}^F \end{pmatrix} \quad (9.5)$$

La matrice de covariance de $\hat{\mu}$ s'en déduit via la formule de propagation des incertitudes :

$$\mathbf{V}_{\mu}^F(\mu) = \mathbf{J}_{\mu} \mathbf{V}_{\text{SN}} \mathbf{J}_{\mu}^T \quad (9.6)$$

où $\mathbf{J}_{\mu} = (\frac{\partial \mu}{\partial X_0}, \frac{\partial \mu}{\partial X_1}, \frac{\partial \mu}{\partial c})$ est la matrice des dérivées de l'équation 9.4.

9.2 Biais

Les résultats présentés dans cette section ont été obtenus à partir d'échantillons d'entraînement LSST-like de taille 500, 1000 et 5000. Nous avons généré 50 réalisations de chacun de ces lots de données.

Pour un lot donné, aucun biais significatif fonction du *redshift* n'est décelable. Les figures 9.1, 9.2 présentent les différences entre les paramètres reconstruits et les paramètres vrais pour m_{B^*} , X_1 , c et μ en fonction du *redshift*, pour des échantillons de 500 et 1000 supernovæ respectivement. Pour chacun des paramètres, une droite est ajustée sur les résidus, afin de quantifier la significativité d'un biais éventuel. Aucun effet significatif n'est mis en évidence.

Les figures 9.3, 9.5 et 9.6 présentent les mêmes tests mais cette fois-ci en moyennant 50 simulations d'échantillons de 500, 1000 et 5000 supernovæ respectivement. On constate que les paramètres X_1 et c sont parfaitement reconstruits, i.e. aucun biais significatif dépendant du *redshift* n'est décelable. (Nous n'observons pas de biais additif car nous nous avons explicitement recentré les distributions des X_1 et c à la génération).

En revanche, toujours en moyennant 50 réalisations, nous observons un faible biais additif sur les X_0 reconstruits. Ce biais décroît avec la taille de l'échantillon. Il est nul pour 5000 supernovæ. Pour 500 et 1000 supernovæ, on déce par ailleurs un biais dépendant du *redshift* (0.4% sur l'ensemble de la gamme en *redshift*) affectant les X_0 et donc les modules de distance standardisés. L'effet est assez faible et là encore disparaît complètement pour 5000 supernovæ. L'essentiel de ce biais semble lié à la force des pénalités de régularisation implémentée dans l'ajustement. Nous avons mené une série de tests additionnels, dans lesquels l'intensité de la régularisation μ_{reg} était divisée par un facteur 10 lors de l'ajustement principal (3e fit). Le résultat est présenté figure 9.4. Le test porte sur 40 réalisations d'un lot d'entraînement LSST-like de 500 supernovæ. Avec une régularisation correctement ajustée, plus aucun biais n'est détectable.

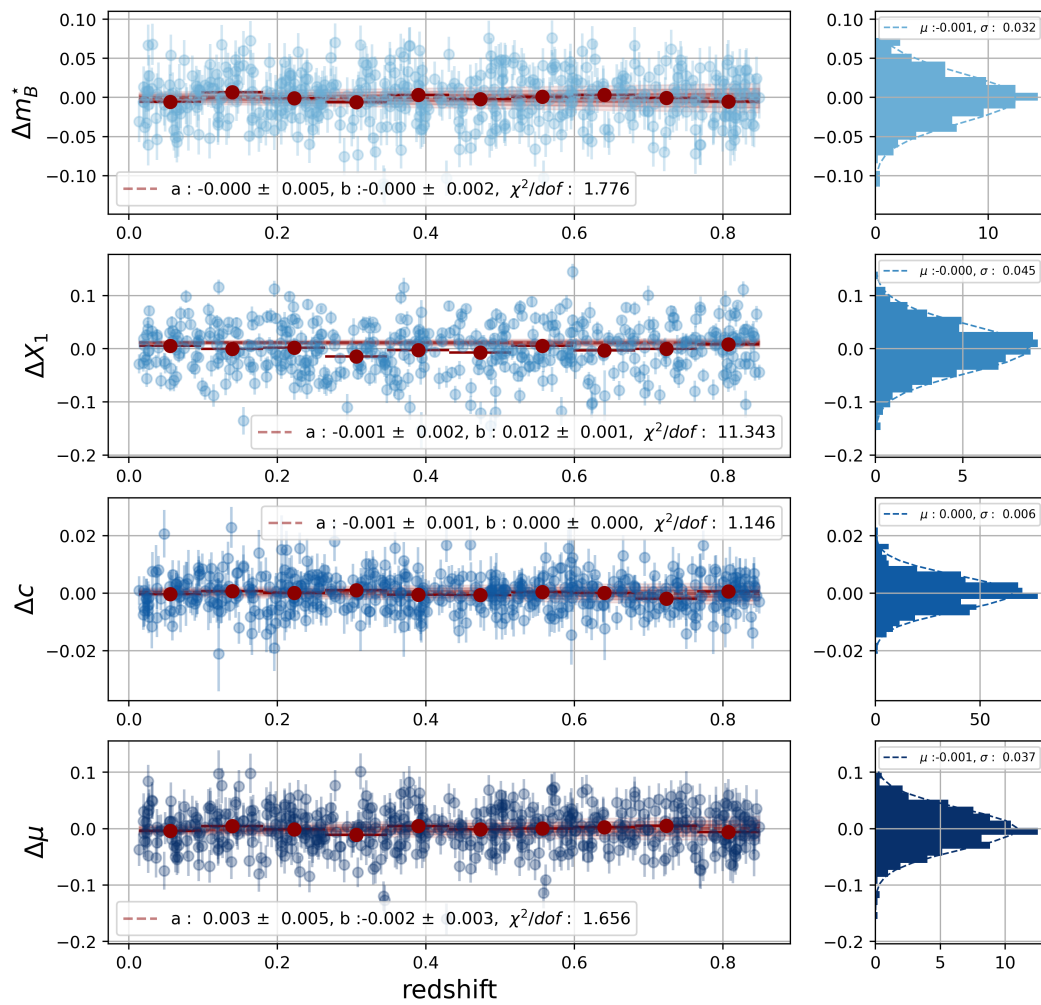


FIGURE 9.1 – Reconstruction des biais sur les paramètres des SNe pour une simulation LSST de 500 supernovæ.

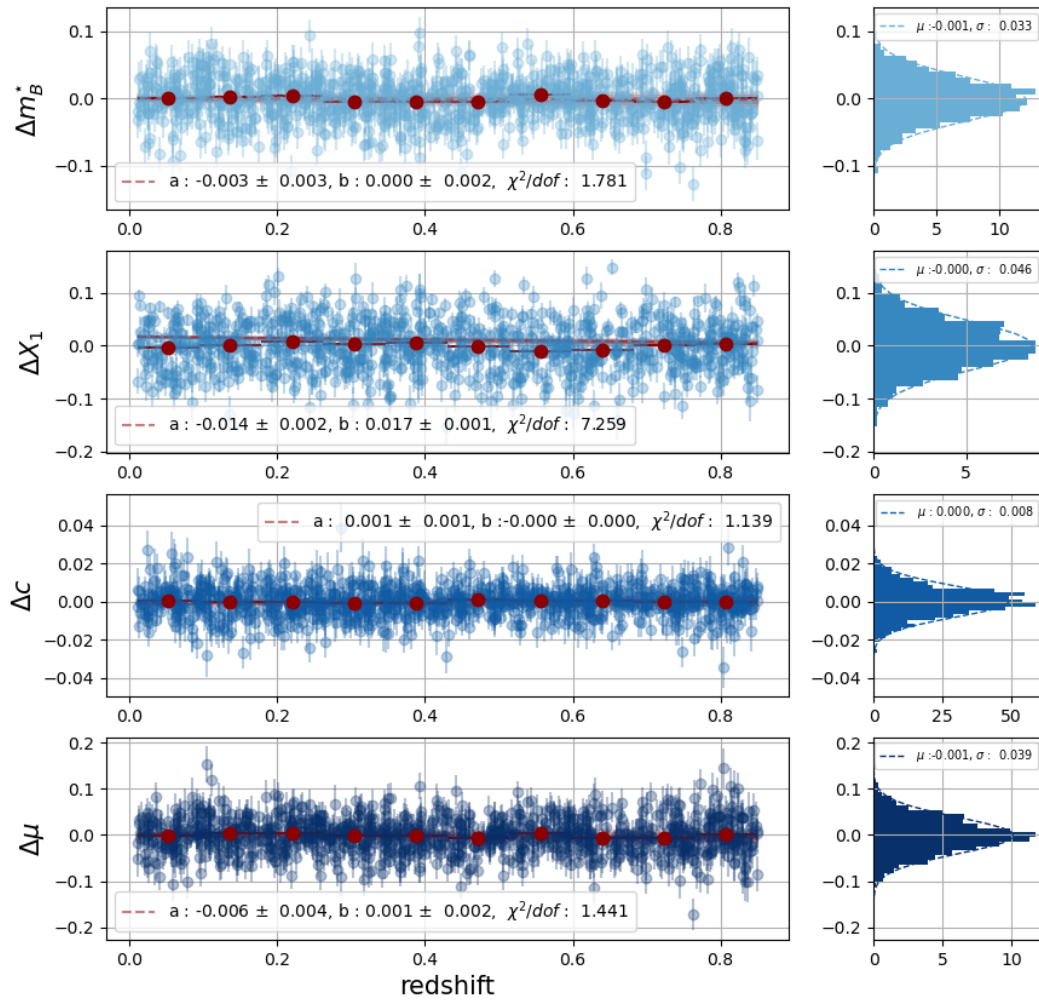


FIGURE 9.2 – Reconstruction des biais sur les paramètres des SNe pour une simulation LSST de 1000 supernovæ.

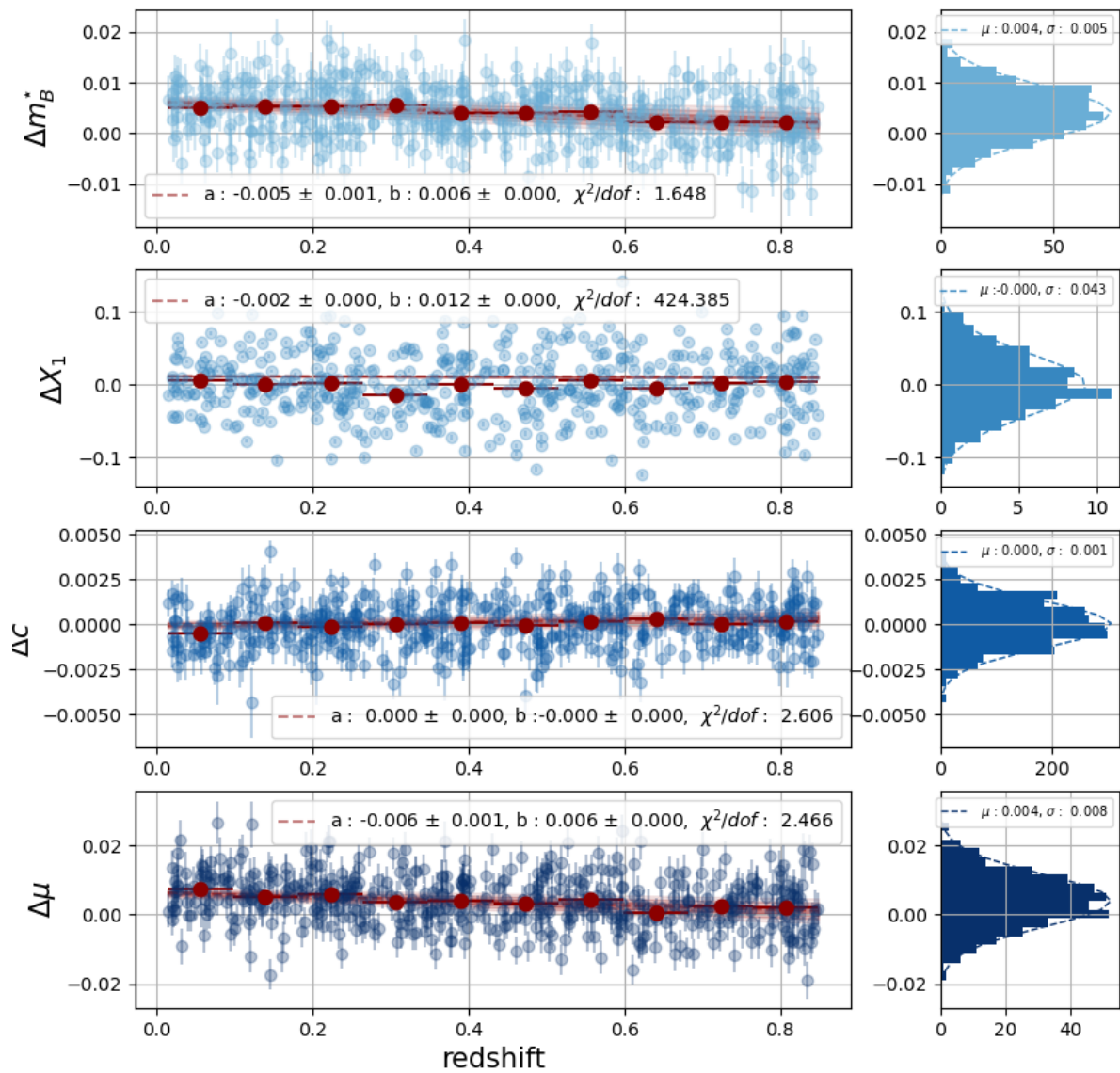


FIGURE 9.3 – Reconstruction des biais sur les paramètres des SNe pour les 50 simulations LSST de 500 supernovæ.

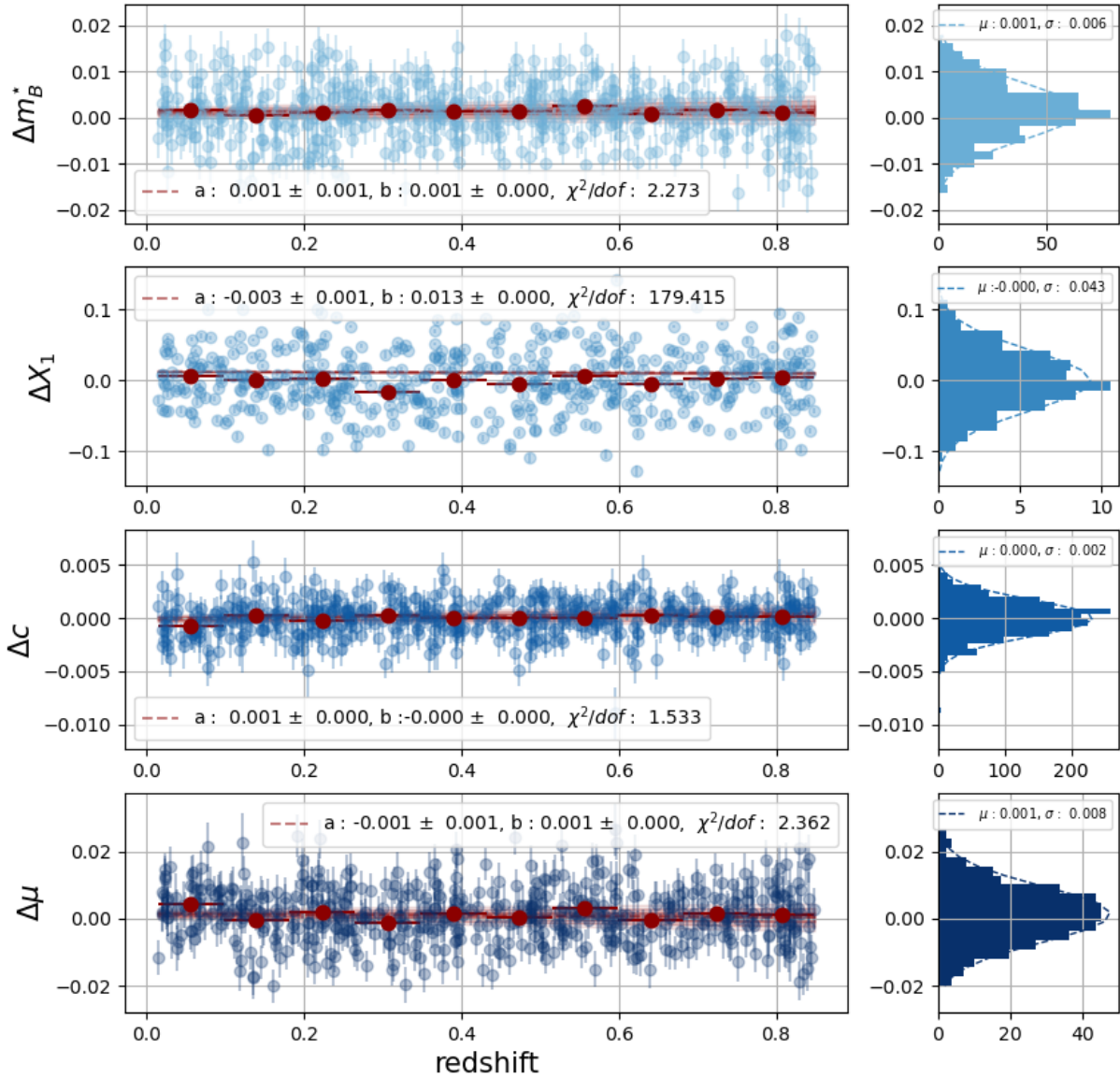


FIGURE 9.4 – Reconstruction des biais sur les paramètres des SNe pour 40 simulations LSST de 500 supernovæ, pour lequel le paramètre de régularisation à été divisé par 10 lors du troisième ajustement.

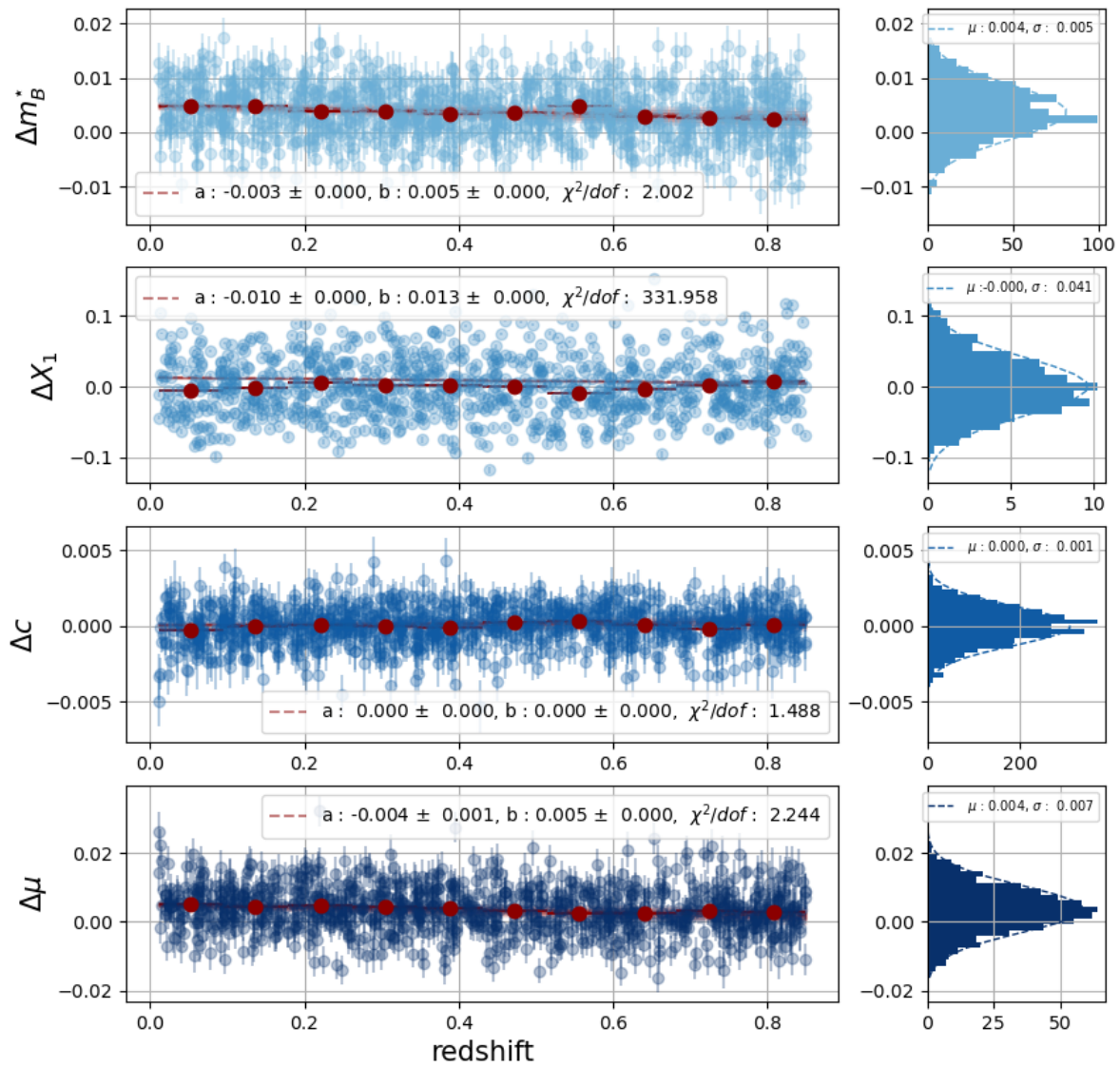


FIGURE 9.5 – Reconstruction des biais sur les paramètres des SNe pour 50 simulations LSST de 1000 supernovæ.

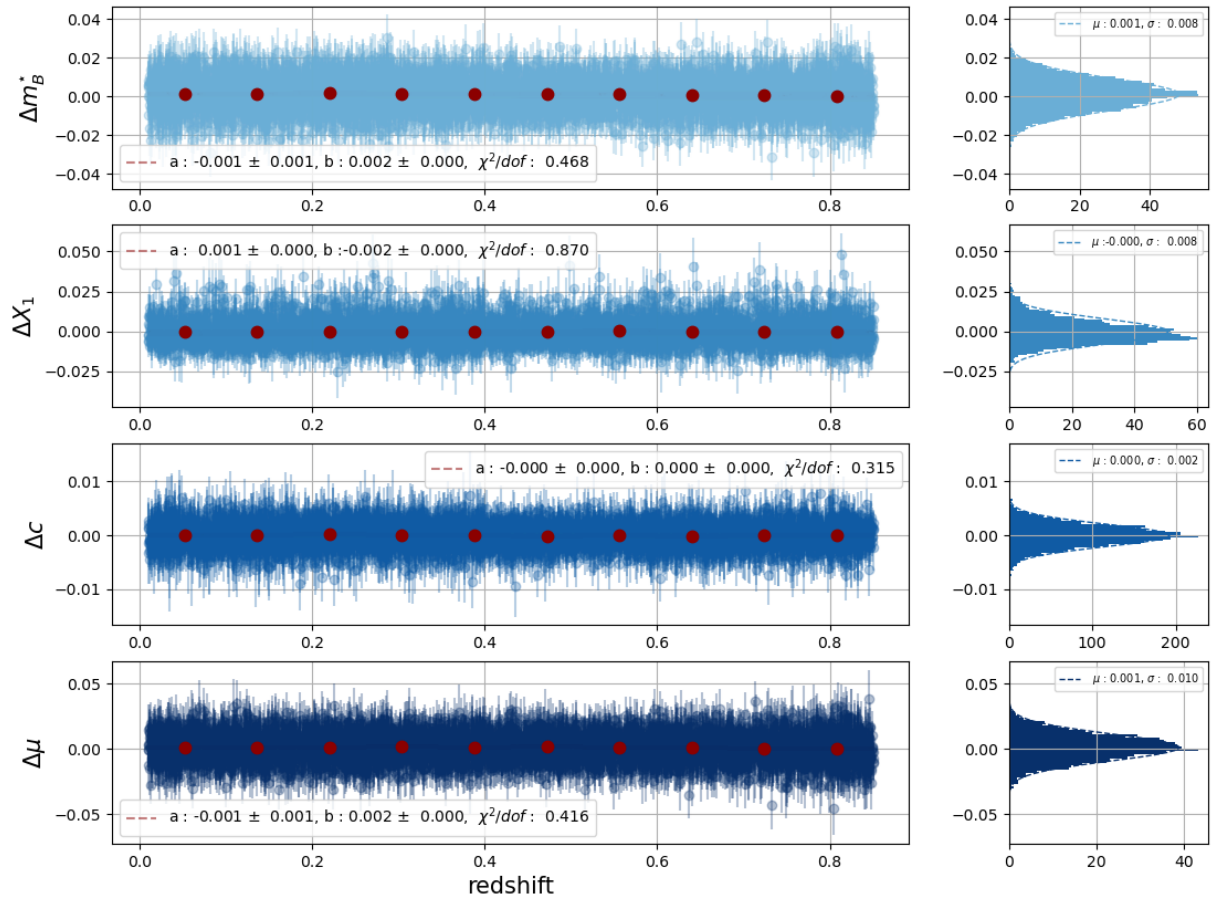


FIGURE 9.6 – Reconstruction des biais sur les paramètres des SNe pour 50 simulations LSST de 5000 supernovæ.

9.3 Variance

Nous passons maintenant aux tests de la variance. La matrice de covariance empirique des paramètres reconstruits est calculée selon la formule suivante :

$$\mathbf{V}_{\mathbf{X}}^{\mathbf{E}} = \frac{(\mathbf{X} - \bar{\mathbf{X}})^T (\mathbf{X} - \bar{\mathbf{X}})}{N_{simu} - 1} \quad (9.7)$$

\mathbf{X} est la matrice des paramètres estimés sur chacun des lots. Chaque colonne correspond aux valeurs reconstruites pour une simulation donnée. Si nous avons simulé 50 fois 500 SNe alors elle est de taille 500×50 . On utilise la valeur moyenne du paramètre \mathbf{X} , $\bar{\mathbf{X}}$ pour se prévenir d'un éventuel biais additif. L'utilisation de la moyenne empirique implique le facteur $\frac{1}{N-1}$ pour compenser la perte d'un degré de liberté.

La matrice de covariance empirique des modules de distance est notée $\mathbf{V}_{\mu}^{\mathbf{E}}$ et calculée comme suit :

$$\mathbf{V}_{\mu}^{\mathbf{E}} = \frac{(\mu - \bar{\mu})^T (\mu - \bar{\mu})}{N_{simu} - 3} \quad (9.8)$$

Cette fois-ci, l'utilisation du module de distance empirique moyen $\bar{\mu}$ requiert de corriger de la perte de 3 degrés de liberté, pour X_0 , c et X_1 . Cette matrice de covariance est comparée avec celle issue de l'ajustement.

9.3.1 Tests sur lots de données LSST-like

Tout d'abord nous nous assurons que la matrice de Fisher de l'ajustement est quasi-indépendante de la réalisation. Cela serait exactement le cas si le modèle était linéaire, ses dérivées étant indépendantes de la position. Ce n'est pas tout à fait vérifié pour la matrice de Fisher d'un ajustement non-linéaire : le minimum du log-vraisemblance varie d'une réalisation à l'autre, et les dérivées aussi. La figure 9.7 montre que les variations relatives des termes diagonaux de la matrice de Fisher n'excèdent pas quelques pourcents (1σ) de réalisation à réalisation. Dans ce qui suit, nous utiliserons pour nos comparaisons l'une des matrices de Fisher issues de la simulation.

Covariance des estimateurs de \hat{X}_1 et \hat{c} Le panel de gauche de la figure 9.8 (resp. 9.9) présente la comparaison entre la matrice de covariance des paramètres X_1 (resp. c) issue du fit (haut) et empirique (bas). Avec 50 réalisations seulement, la matrice de covariance empirique est plus bruitée. On distingue néanmoins que les deux matrices ont la même structure.

Plus quantitativement, sur les panneaux du milieu, on représente la racine carrée de la diagonale de chacune de ces matrices de covariance (empirique : points gris, Fisher : triangles bleus) en fonction du *redshift* (haut) ainsi que les ratios de ces deux diagonales (bas). On constate que la matrice de covariance issue de l'ajustement décrit correctement les covariance empiriques. Cela signifie que les incertitudes de mesure, de calibration et de modèle sont bien propagées jusqu'aux paramètres X_1 et c . Le panel de droite montre les histogrammes correspondants.

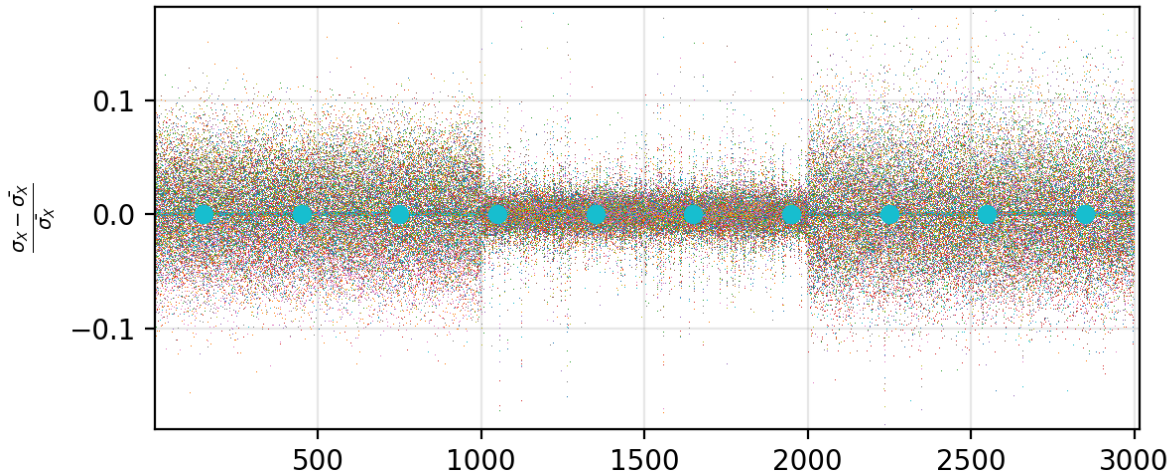


FIGURE 9.7 – Comparaison de incertitudes sur les paramètres d'intérêt (X_0 : 0 à 1000, X_1 : 1 à 2000 et c : 2000 à 3000) extraites des matrices de Fisher obtenues pour les 50 ajustements.

Covariance de \hat{X}_0 et $\hat{\mu}$ Les figures 9.10 et 9.11 présentent les mêmes quantités pour les estimateurs de m_{B^*} et μ . Pour ces quantités, on observe que la variance de Fisher est affectée d'un biais de l'ordre de 0.767 ± 0.085 . La variance prédite par l'ajustement est inférieure de 25% à la variance empirique.

Une investigation détaillée a permis d'établir que là encore, le problème était lié à notre modélisation du *color scatter*. Nous avons vu au chapitre précédent que la variance des paramètres κ reconstruits était inférieure à la variance générée. En pratique, c'est le terme constant du prior $\sigma_\kappa(\lambda)$ qui est biaisé. Ce biais impacte la variance de Fisher des paramètres X_0 , et de là, la variance des modules de distance standardisés.

Pour clarifier ce point, nous avons réalisé une série d'entraînements, dans lesquels le terme constant du polynôme $\sigma_\kappa(\lambda)$ était fixé à la valeur générée. Le résultat est présenté sur les figures 9.12 et 9.13. Il montre que la correction de ce simple biais permet d'obtenir l'estimation correcte de l'incertitude sur \hat{X}_0 et $\hat{\mu}$.

Une amélioration de la modélisation du *color scatter* est nécessaire pour améliorer ce point. Comme discuté au chapitre précédent, la solution semble être de diminuer le nombre de variables latentes, par exemple, passer de 1 paramètre κ par bande à une fonction *smooth* de deux ou trois paramètres par SN, cette fonction étant fixée à 1 aux alentours de λ_{B^*} . Cela permettrait de briser les interactions entre *color scatter* et X_0 et d'éviter que le biais de variance se propage sur l'estimateur de distance. Si un tel biais subsistait, la solution consistera à le modéliser et à le corriger au moment du calcul de la matrice de Fisher finale.

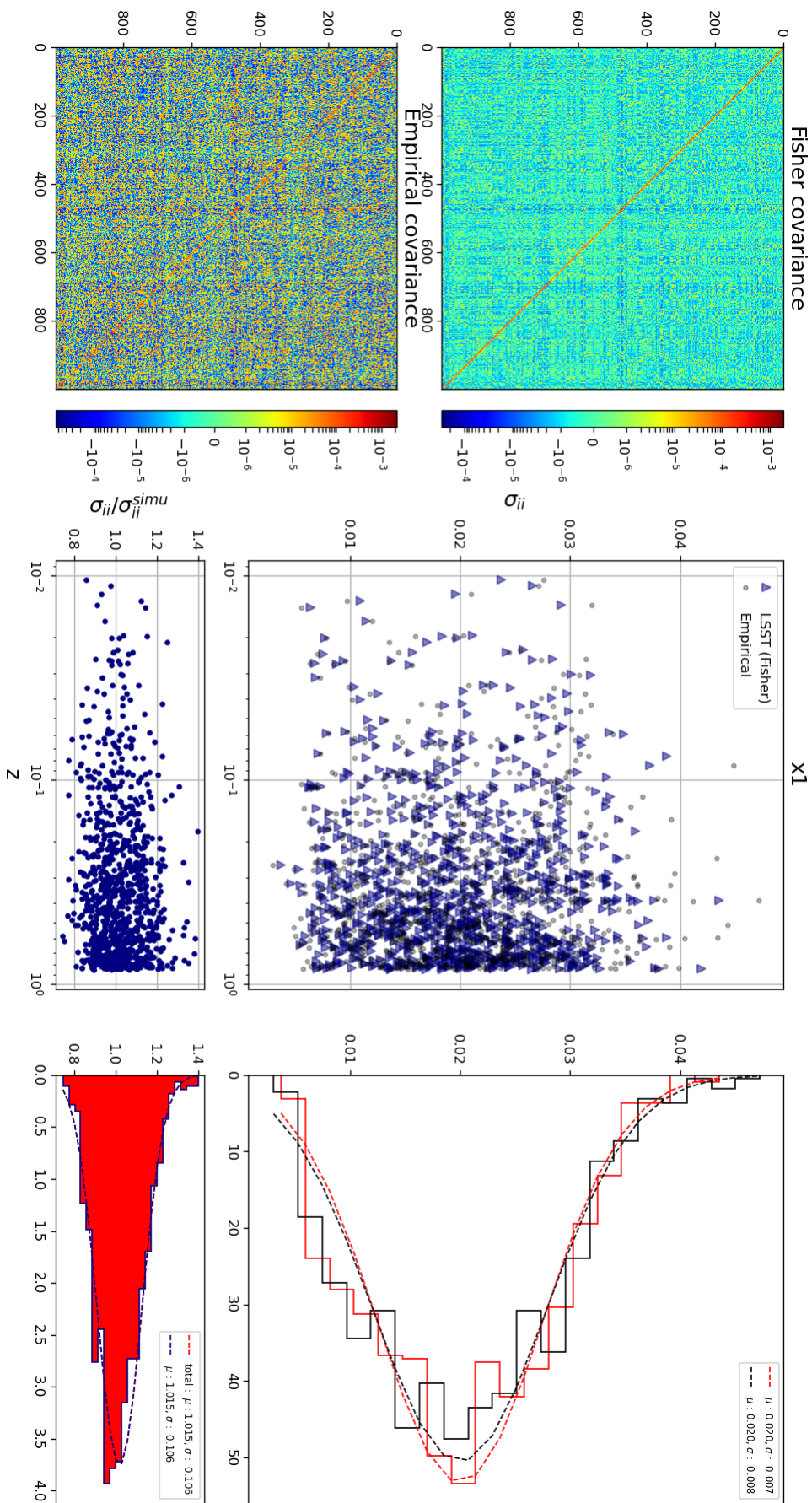


FIGURE 9.8 – Test de couverture pour les couleurs X_1 , pour l'ajustement du modèle sur les simulations LSST de 1000 SNe.

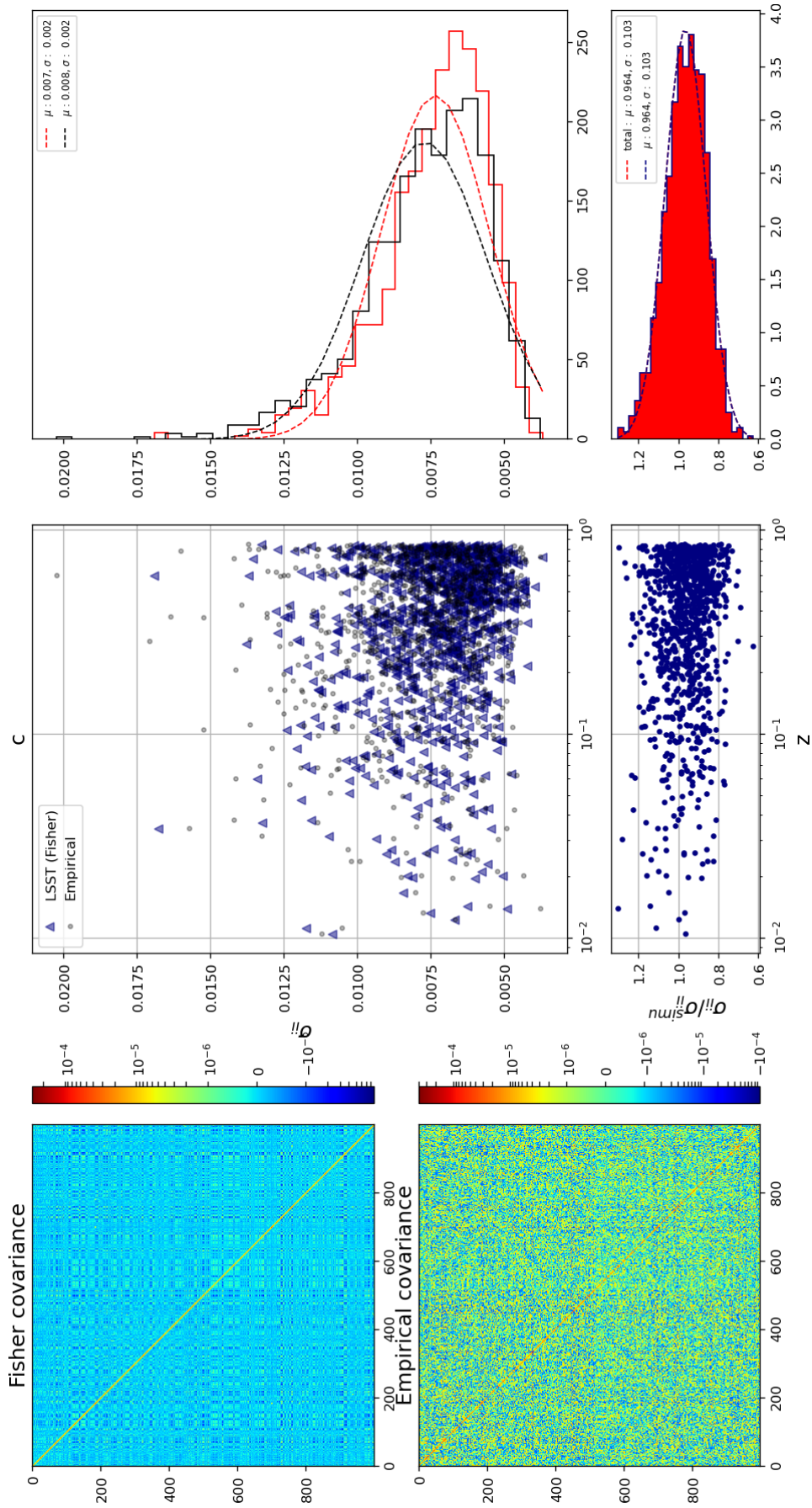


FIGURE 9.9 – Test de couverture pour les paramètres de forme c , pour l’ajustement du modèle sur les simulations LSST de 1000 SNe.

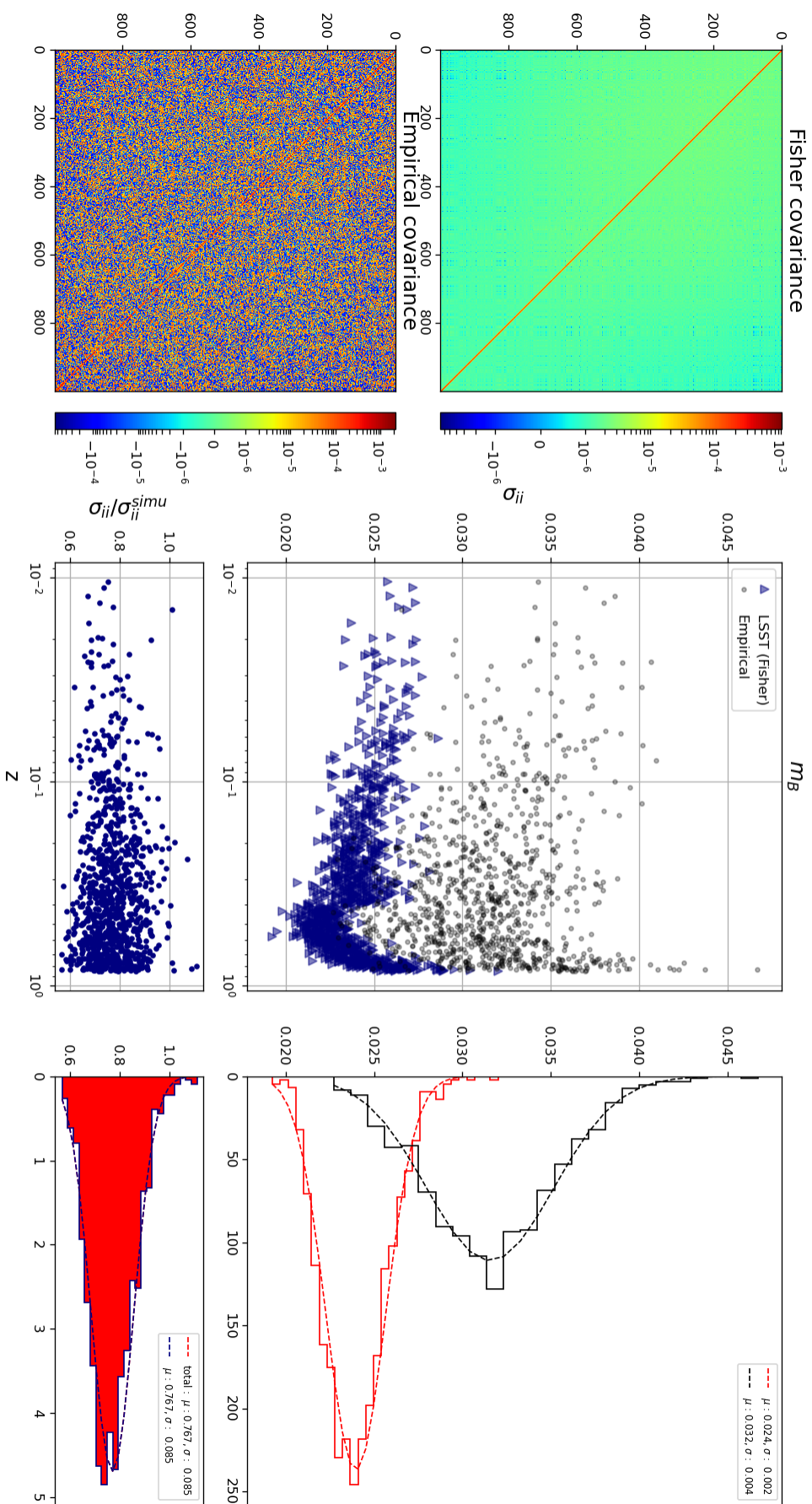


FIGURE 9.10 – Test de couverture pour la magnitude au maximum de luminosité en bande B^* , pour l'ajustement du modèle sur les simulations LSST de 1000 SNe.

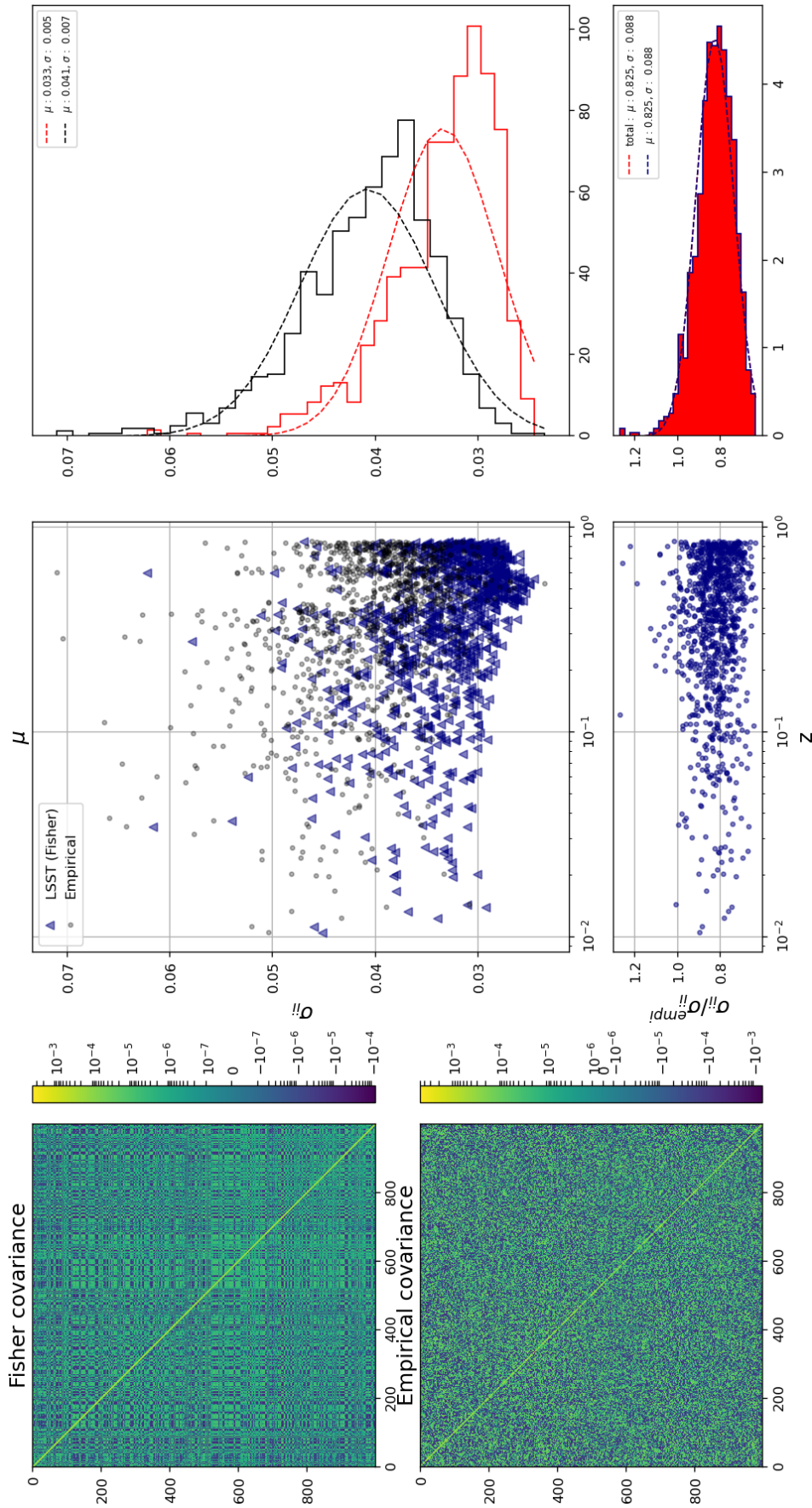


FIGURE 9.11 – Test de couverture pour le module de distance μ , pour l’ajustement du modèle sur les simulations LSST de 1000 SNe.

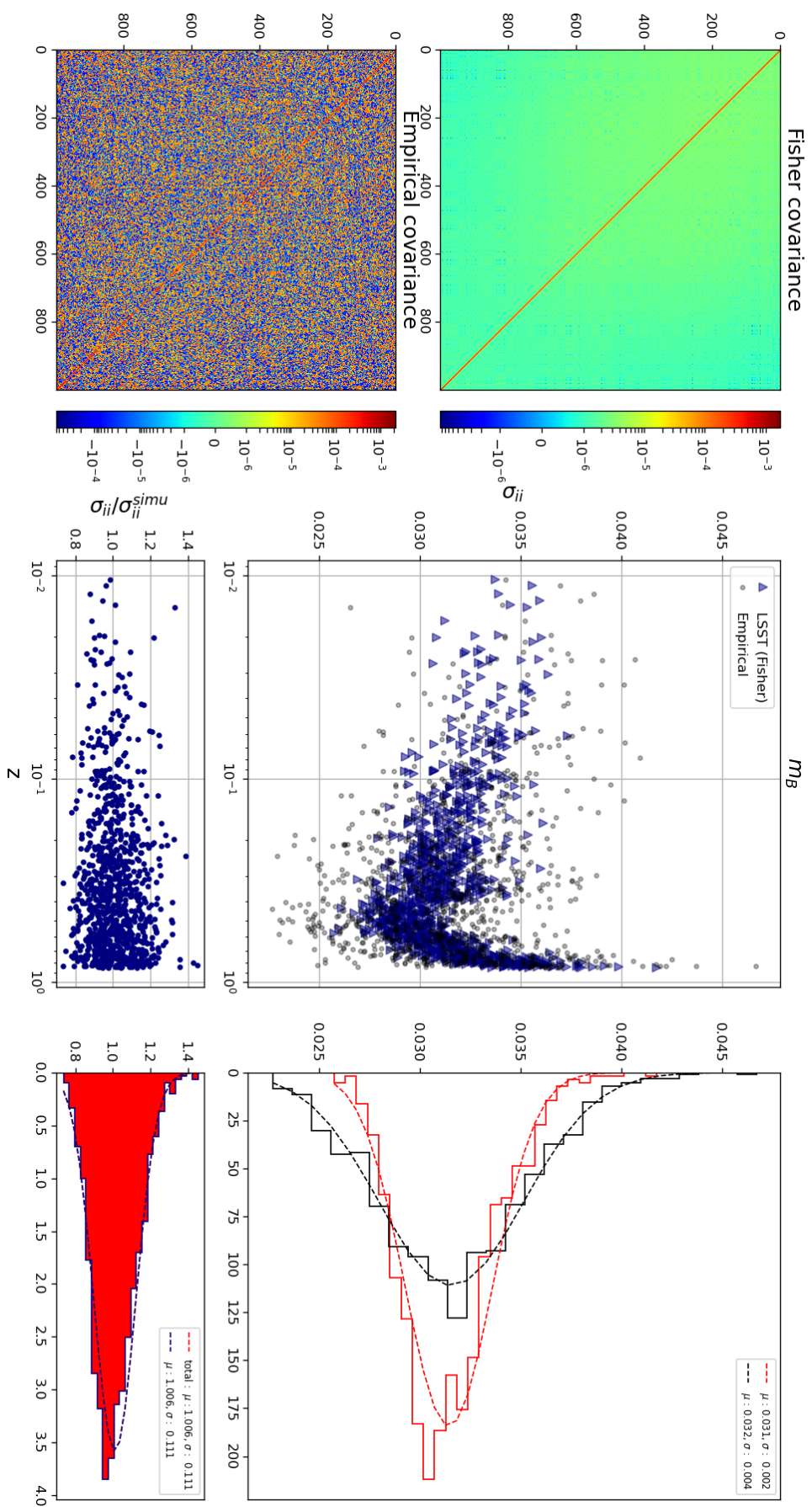


FIGURE 9.12 – Test de couverture pour la magnitude au maximum de luminosité en bande B^* , pour l'ajustement d'un modèle sur les simulations LSST de 1000 SNe, calculé avec une matrice de Fisher pour laquelle le terme constant de la dispersion en couleur est dé-biaisé.

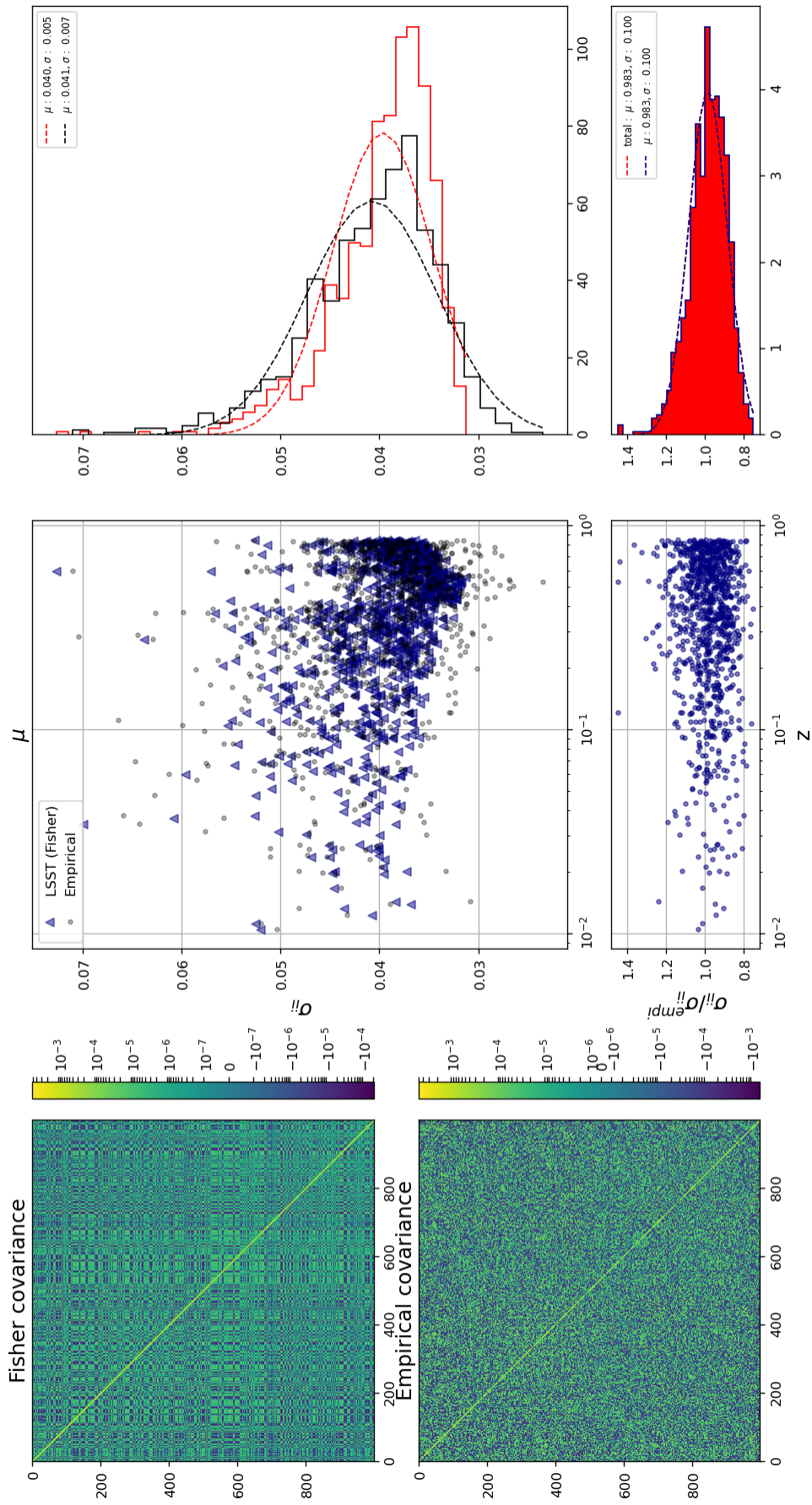


FIGURE 9.13 – Test de couverture pour le module de distance constant μ , pour l’ajustement d’un modèle sur les simulations LSST de 1000 SNe, calculé avec une matrice de Fisher pour laquelle le terme constant de la dispersion en couleur est dé-biaisé.

9.3.2 JLA

Enfin, nous avons effectué le même type de tests sur le lot de données (simulées) JLA. Ces tests ont été menés sans génération ni modélisation du *color scatter*. En effet, nous avons montré au chapitre 8 que la modélisation du *color scatter* actuellement en place ne permettait pas la convergence du modèle complet.

Nous avons généré 30 réalisations du lot JLA, en activant les sources de bruit suivantes : mesure, *error snake*, calibration. Le bruit de calibration est généré suivant la diagonale de la matrice JLA.

Biais La comparaison entre paramètres d'intérêt générés et reconstruits pour une unique simulation est montrée figure 9.14. Tous les paramètres sont correctement reconstruits et aucun biais significatif dépendant du *redshift* n'est détectable.

En moyennant trente réalisations, en revanche, nous commençons à détecter un biais significatif dépendant du *redshift*, (figure 9.15, droite). Ce biais semble extrêmement sensible à la précision de la calibration. En réalisant la même étude sur 30 réalisations avec une précision de calibration de 0.1%, le biais disparaît entièrement (figure 9.15, gauche). Dans le cas de JLA, il semble que la multiplicité d'instruments et de bandes en jeu soit un facteur de fragilité, dans la mesure où chaque instrument peuple une section spécifique du diagramme de Hubble.

Variance Enfin, nous avons, sur ces mêmes simulations, comparé variance de Fisher issue de l'ajustement et variance empirique calculée sur 30 simulations. Les résultats sont présentés sur les figures 9.16, 9.17, 9.18 et 9.19. La modélisation du *color scatter* étant inactive pour ces tests, on constate un parfait accord entre variance de Fisher et variance Monte-Carlo. Le schéma d'entraînement que nous avons mis au point permet de propager de manière exacte, jusqu'aux distances, toutes les incertitudes affectant les données, y compris pour un lot de données hétérogène tel que JLA.

9.4 Conclusion

Dans ce chapitre, nous avons mené plusieurs tests portant sur le biais et la variance de l'estimateur de distance issu de l'entraînement du modèle. Nous n'avons fait qu'effleurer le sujet : la caractérisation de l'estimateur en fonction (1) du bruit de mesure (2) du sampling des courbes de lumière (3) du design du relevé (4) de la précision de la calibration (5) du réglage des hyperparamètres etc. mériterait plusieurs chapitres et plusieurs mois de travail. Nous avons néanmoins établi plusieurs résultats importants.

Tout d'abord, l'estimateur de distance semble non biaisé. À l'échelle d'une simulation, aucun biais fonction du *redshift* n'est détectable sur les échantillons LSST-like. En stackant de l'ordre de cinquante simulations, un biais a été détecté. Il a pu être réduit à zero en réglant l'intensité de la régularisation implémentée dans l'ajustement.

Ensuite, nous avons établi qu'en l'absence de modélisation du *color scatter*, la matrice de covariance issue de l'ajustement reproduit parfaitement la matrice de covariance empirique calculable par Monte-Carlo. Cela signifie que nous sommes capables de propager (à faible coût) jusqu'aux distances la totalité des incertitudes affectant le lot d'entraînement (mesure, calibration, modélisation). La modélisation du *color scatter* doit encore être améliorée. Dans l'état actuel, le biais sur la variance des paramètres κ se propage directement l'estimation de la variance des X_0 et donc des distances. Plusieurs stratégies sont à l'étude pour résoudre ce problème.

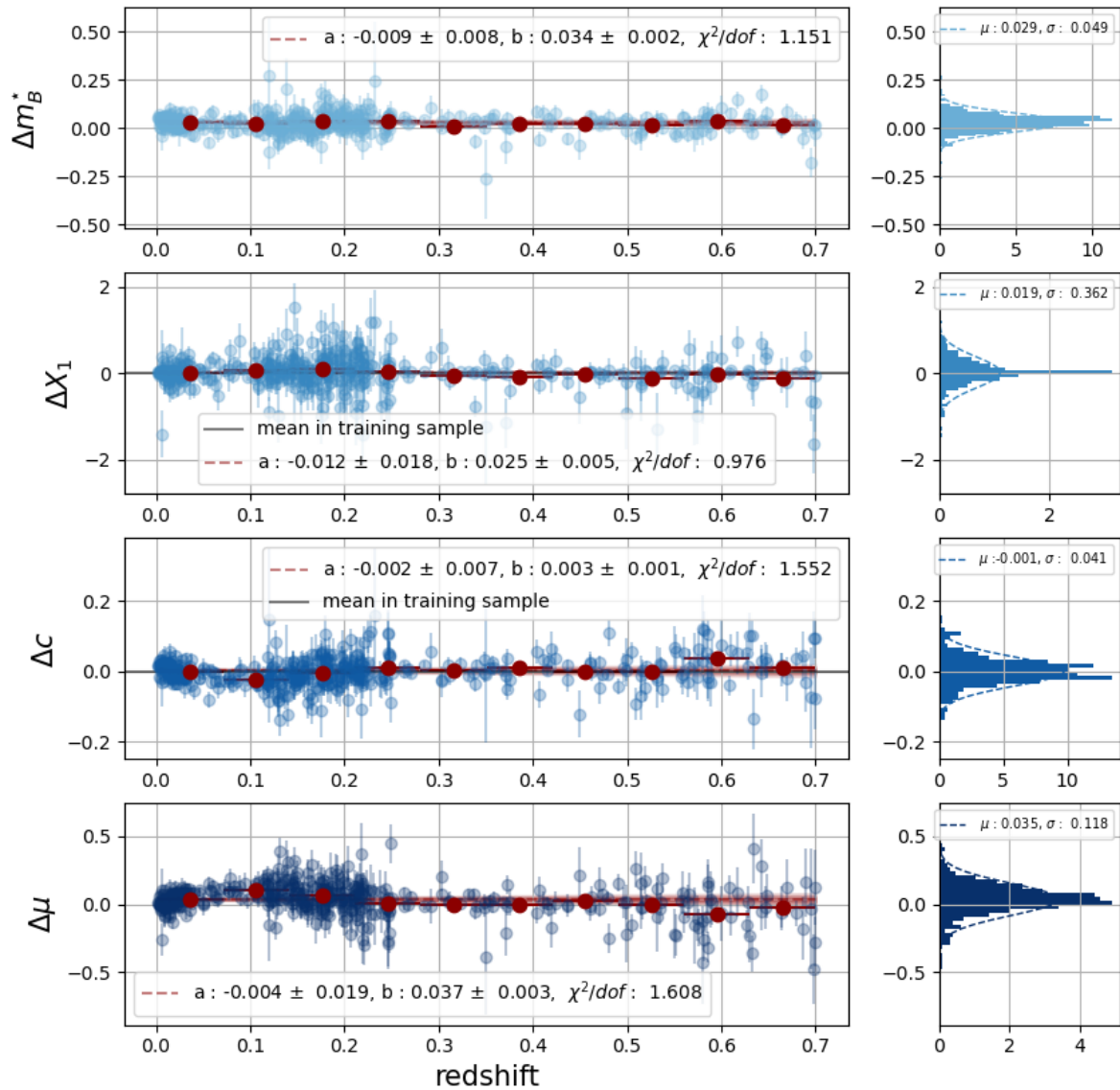


FIGURE 9.14 – Biais sur les paramètres d'intérêts issu de l'ajustement du modèle et de l'*error snake* sur JLA. Pour chaque paramètre, m_{B^*} , X_1 , c sont les erreurs extraites directement de la matrice de Fisher.

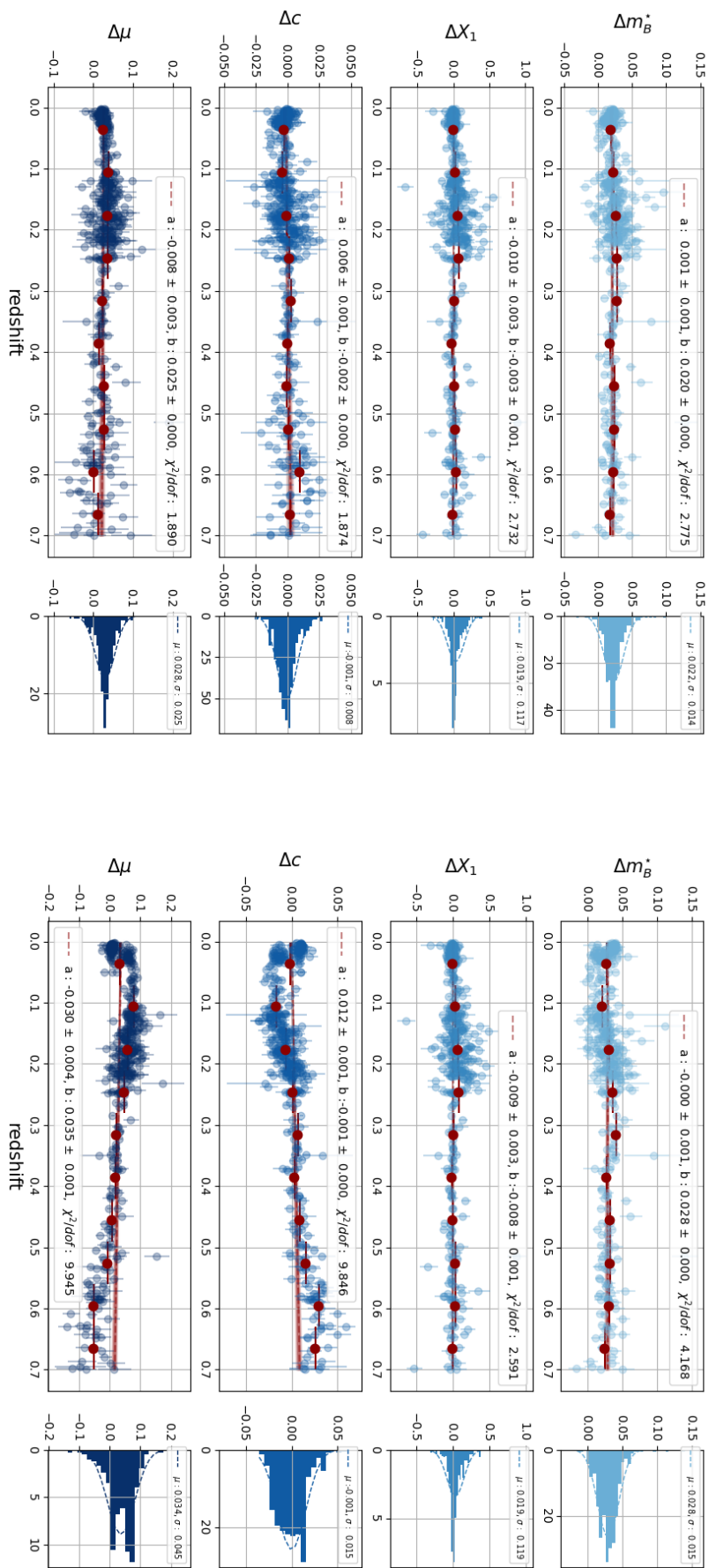


FIGURE 9.15 – Biais sur les paramètres d'intérêts issu de l'ajustement du modèle et de l'*error snake* sur JLA. À gauche pour l'ensemble des simulations pour lesquelles le prior de calibration est tenu au pour mille, à droite quand il est donné par la matrice de calibration de JLA.

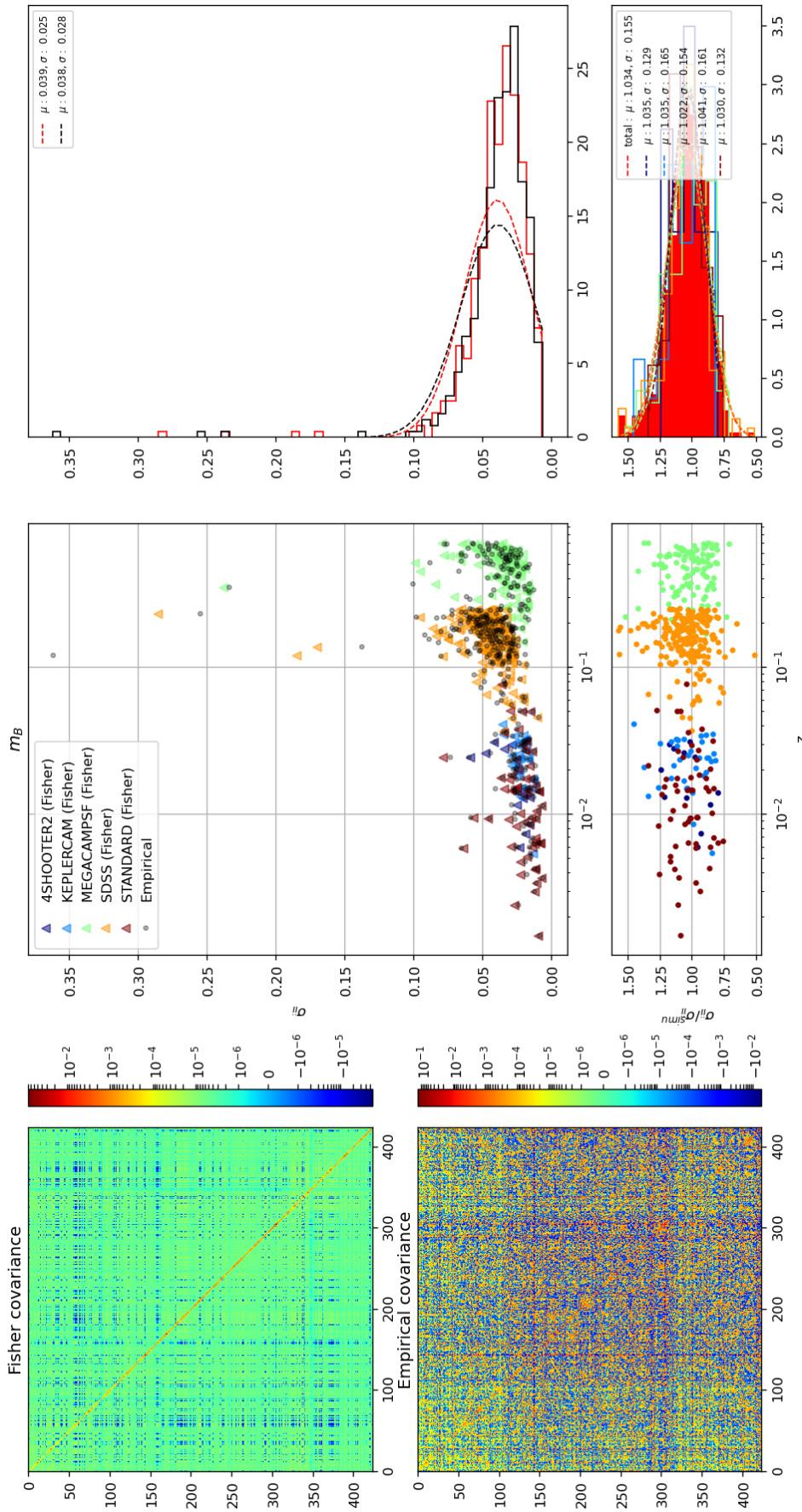


FIGURE 9.16 – Test de couverture pour la magnitude au maximum de luminosité en bande B^* , pour l’ajustement d’un modèle propageant les incertitudes de calibration de JLA sur le lot d’entraînement de JLA. En haut la matrice $V^F(\mu)$ définie équation 9.6), en dessous la matrice de covariance empirique $V^E(\mu)$ (cf. équation 9.8). Sur le panneau supérieur du milieu, on peut voir les déviations standard sur les magnitudes, la racine carré de la diagonale des matrices de covariances en fonction du milieu, les déviations standard de Fisher issu de la hessien de l’ajustement sont représentées par des triangles, celles empiriques sont représentées par des triangles noirs. Le code couleur code les différents relevés de supernovæ composant JLA. En dessous, sont présentés les rapports de ces estimations, dont la valeur moyenne est 1. À droite, sont affichées les histogrammes correspondant.

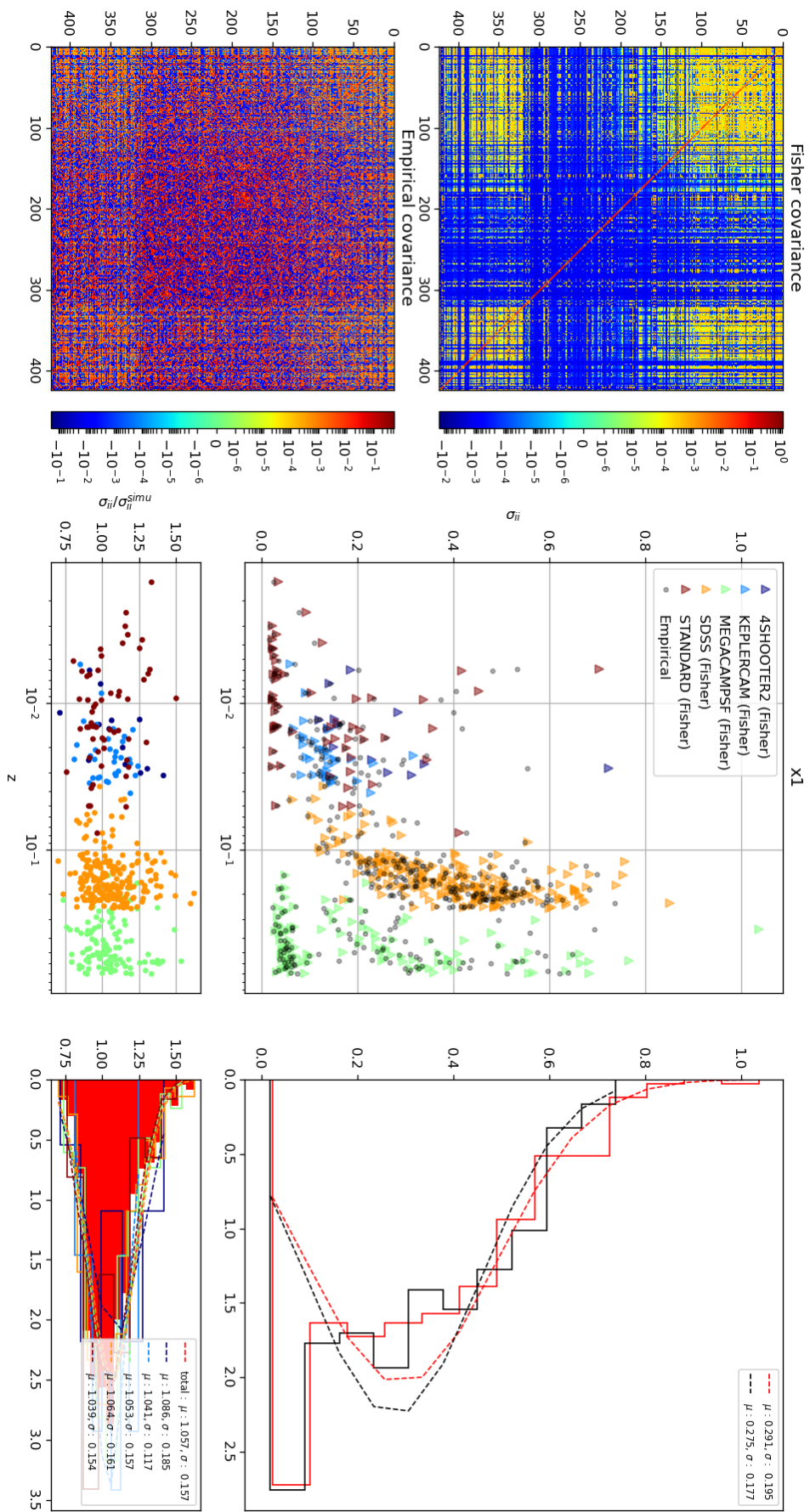


FIGURE 9.17 – Test de couverture pour les paramètres de forme X_1 , pour l'ajustement d'un modèle propageant les incertitudes de calibration de JLA sur le lot d'entraînement de JLA. Le ratio des déviations standards entre les triangles colorés (issu de la matrice de Fisher) et les déviations standards empiriques, en noirs a une valeur moyenne est 1.

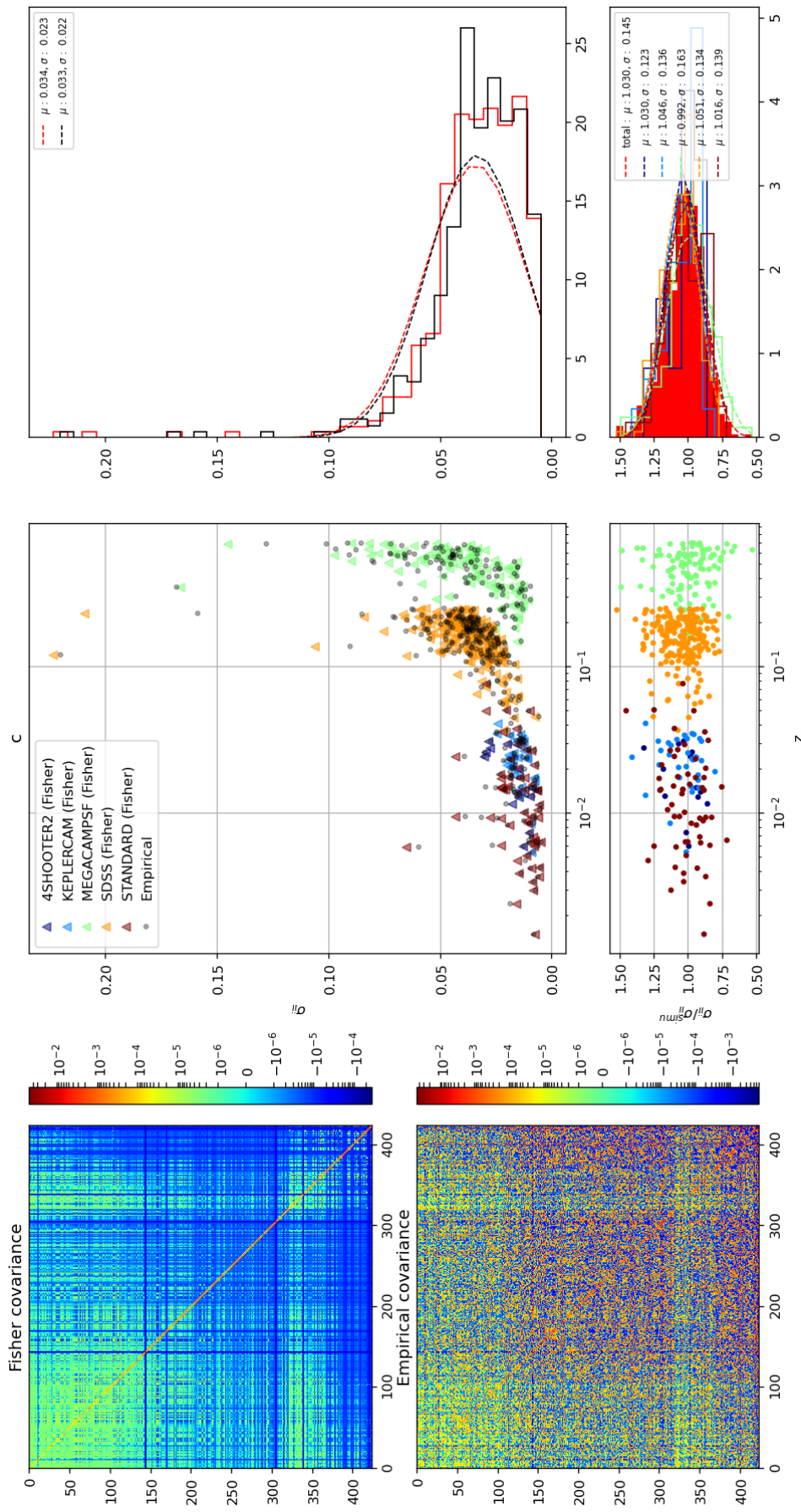


FIGURE 9.18 – Test de couverture pour les paramètres de couleur c , pour l’ajustement d’un modèle propageant les incertitudes de calibration de JLA sur le lot d’entraînement de JLA. Le ratio des déviations standards entre les triangles standards (issu de la matrice de Fisher) et les déviations standards empiriques, en noirs a une valeur moyenne est 1.

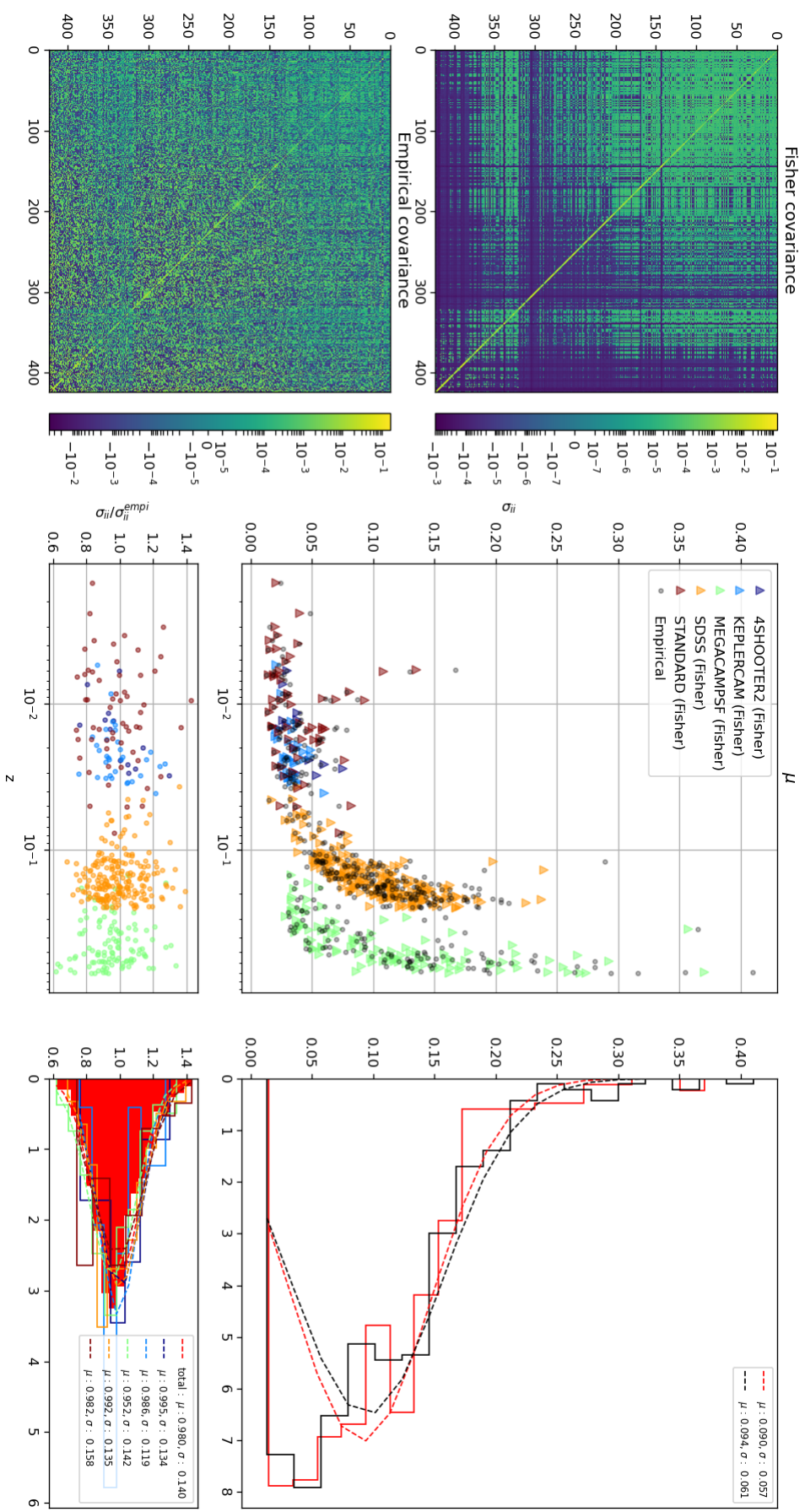


FIGURE 9.19 – Test de couverture pour les modules de distance μ , pour l'ajustement d'un modèle propageant les incertitudes de calibration de JLA sur le lot d'entraînement de JLA. Le rapport des déviations standards entre les triangles colorés (issu de la matrice de Fisher) et les déviations standards empiriques, en noirs a une valeur moyenne est 1, à 13% près.

Conclusion

Durant cette thèse j’ai développé un ensemble d’algorithmes comprenant une modélisation empirique spectrophotométrique de supernovæ de type Ia, un minimiseur permettant d’ajuster simultanément les paramètres du modèle et de son modèle d’erreur et un *pipeline* de simulations. L’ensemble est nommé NaCl pour *Nouveaux algorithmes de Courbes de lumière*.

Le modèle est entraîné simultanément sur des courbes de lumière et des spectres. Il suit la paramétrisation de SALT2 et permet de décrire la diversité des supernovæ de type Ia avec deux paramètres : un paramètre de forme X_1 et un paramètre de couleur c (cf. chapitre 5). Pour décrire la variabilité résiduelle non capturée par le modèle, j’ai également développé un modèle d’erreur. Celui-ci comprend une composante diagonale modélisant la variance additionnelle observée dans les données et non réductible au bruit de mesure, et une composante non diagonale qui décrit la dispersion cohérente observée dans les données autour de la loi de couleur moyenne du modèle. De plus, j’ai étendu le modèle pour permettre de propager directement dans l’entraînement les incertitudes de calibration photométriques. Enfin, je me suis assuré que la date de maximum de luminosité pourrait être ajustée simultanément avec tous les autres paramètres.

Pour entraîner le modèle, j’ai développé un minimiseur dont les détails sont décrits au chapitre 6. Il permet un entraînement rapide et simultané du modèle et du modèle d’erreur. Il n’était pas a priori évident que cet ajustement puisse converger. Un résultat important de cette thèse est qu’il est possible de réaliser l’entraînement d’un modèle de type SALT2 en une minimisation unique.

Cette procédure permet une propagation directe et exacte des incertitudes de mesure, de calibration et de modélisation sur tous les paramètres du modèle. Pour permettre l’étude de nouveau(x) paramètre(s) de standardisation, grâce à des modèles étendus, nous avons veillé à ce que l’architecture du minimiseur soit complètement distincte de la paramétrisation du modèle. Ainsi avec l’avènement des nouveaux lots d’entraînement de grande statistique, composés d’observations de DES, HSC, ZTF ou encore dans un futur plus lointain LSST, le code pourra rapidement permettre l’exploration de nouvelles paramétrisations.

Pour faciliter le développement du modèle et du minimiseur et pour étudier les biais de cette procédure de minimisation, j’ai développé plusieurs jeux de simulations, toutes décrites au chapitre 7. D’abord, de simples simulations ont été utilisées pour développer les outils décrits précédemment. J’ai ensuite développé des simulations idéales du relevé LSST, avec des erreurs de mesure photométriques réalistes et dépendantes du *redshift*. Les cadences d’observations ont été choisies pour obtenir des données très bien échantillonnées mais peuvent être dégradées pour reproduire des observations plus réalistes. Ensuite j’ai réalisé des simulations des lots d’entraînement de

JLA et K21 afin de tester la procédure sur des données avec un échantillonnage réel.

Au chapitre 8, j'introduis la procédure de minimisation développée. Elle permet de reconstruire le modèle et le modèle d'erreur simultanément après deux étapes d'initialisation. Dans ce chapitre l'étude du temps de convergence de l'ajustement montre que nous pouvons entraîner notre modèle sur des lots denses de quelques milliers de SNe en quelques heures. Pour les lots simulés JLA et K21, l'ajustement converge en une demi heure et une heure respectivement. Des études préliminaires montrent qu'il sera probablement possible d'accélérer l'entraînement d'un facteur 2 à 5 par rapport à la situation actuelle, sans modification majeure du code, en ayant recours au *multithreading* et/ou en utilisant les ressources GPU à notre disposition.

L'assemblage des tous ces outils, l'ajustement du modèle sur les simulations, m'a permis de construire un *pipeline* pour tester les biais et l'efficacité de l'estimateur de distances issues de l'ajustement (cf. chapitre 9). Grâce aux simulations rapides du relevé LSST, j'ai pu montrer d'abord que les distances reconstruites étaient asymptotiquement non biaisées. Ensuite, j'ai établi que, sans modélisation du *color scatter*, l'estimateur était efficace : sa variance atteint la borne de Cramer-Rao. La variance issue du fit décrit correctement la variance Monte-Carlo. Cette variance totale inclut le bruit de mesure, les incertitudes de calibration et de modélisation. Cela représente une simplification considérable par rapport aux procédures de propagation des incertitudes mises en œuvre jusqu'à présent dans le domaine.

Un secteur du modèle nécessite encore quelques développements : la modélisation du *color scatter*. Notre implémentation actuelle introduit trop de degrés de liberté dans le modèle, ce qui fait diverger certains entraînements lorsque la statistique de l'échantillon est trop faible. Par ailleurs, elle est la source probable d'un biais de $\sim 25\%$ sur la variance de l'estimateur de distance. Plusieurs solutions sont à l'étude pour améliorer ce point.

Le code développé au cours de cette thèse permet de ne plus dissocier les étapes d'"entraînement" et d'"ajustement des courbes de lumière". Le code est suffisamment rapide pour réaliser les deux opérations en une seule minimisation et fournir un ensemble de distances, accompagnées de leur matrice de covariance. À plus long terme, nous souhaitons tirer parti des lots de très grande qualité en cours de publication (Pantheon, DES, ZTF, SNfactory) pour étendre la paramétrisation du modèle, affiner la description des SNe Ia et améliorer la qualité de l'estimateur de distance. À terme, il est probablement possible de combiner entraînement, modélisation des biais de sélection et ajustement cosmologique en un unique estimateur. L'horizon pour un tel travail est le fantastique lot de données qui sera produit par LSST.

Annexe A

Dérivées

Dans cette partie, nous dérivons le modèle défini équation 5.28 pour la partie spectrale et équation 5.34 pour la partie photométrique. Les dérivées du modèle d'erreur sont explicitées. Nous avons reproduit le modèle d'erreur de l'entraînement de SALT3, nous définissons son algèbre et ses dérivées.

A.1 Modèle

Dérivées de l'équation 5.28 par rapport à tous les paramètres :

$$\frac{\partial S(p, \lambda)}{\partial t_{max}^{B^*}} = \frac{-1}{(1+z)^2} [\theta_i^0 \mathcal{B}'_i(p, \lambda) + X_1 \theta_i^1 \mathcal{B}'_i(p, \lambda)] 10^{0.4cCL(\lambda_r)} s(\lambda_0) \quad (\text{A.1})$$

$$\frac{\partial S(p, \lambda)}{\partial X_1} = \frac{1}{1+z} [\theta_i^1 \mathcal{B}_i(p, \lambda)] 10^{0.4cCL(\lambda_r)} s(\lambda_0) \quad (\text{A.2})$$

$$\frac{\partial S(p, \lambda)}{\partial c} = \frac{1}{1+z} [\theta_i^0 \mathcal{B}_i(p, \lambda) + X_1 \theta_i^1 \mathcal{B}_i(p, \lambda)] 10^{0.4cCL(\lambda_r)} s(\lambda_0) \times 0.4 CL(\lambda_r) \ln(10) \quad (\text{A.3})$$

$$\frac{\partial S(p, \lambda)}{\partial \theta_i^0} = \frac{1}{1+z} \mathcal{B}_i(p, \lambda) 10^{0.4cCL(\lambda_r)} s(\lambda_0) \quad (\text{A.4})$$

$$\frac{\partial S(p, \lambda)}{\partial \theta_i^1} = \frac{1}{1+z} X_1 \mathcal{B}_i(p, \lambda) 10^{0.4cCL(\lambda_r)} s(\lambda_0) \quad (\text{A.5})$$

$$\frac{dS(p, \lambda)}{dp_i} = \frac{1}{1+z} [\theta_i^0 \mathcal{B}_i(p, \lambda) + X_1 \theta_i^1 \mathcal{B}_i(p, \lambda)] 10^{0.4cCL(\lambda)} s(\lambda_0) \times 0.4 c \ln(10) \frac{dCL(\lambda_r)}{dp_i} \quad (\text{A.6})$$

$$\frac{\partial CL(\lambda_r)}{\partial p_i} = \left\{ \begin{array}{ll} \lambda_r^{(5-i)} - \lambda_r & \text{si } \lambda_{UV} < \lambda_r < \lambda_{IR} \\ \left[((5-i) \lambda_{UV}^{(5-i)-1} - 1) (\lambda_r - \lambda_{UV}) + (\lambda_{UV}^{N-i} - \lambda_{UV}) \right] & \text{si } \lambda_r < \lambda_{UV} \\ \left[((5-i) \lambda_{IR}^{(5-i)-1} - 1) (\lambda_r - \lambda_{IR}) + (\lambda_{IR}^{N-i} - \lambda_{IR}) \right] & \text{si } \lambda_{IR} < \lambda_r \end{array} \right\} \quad (\text{A.7})$$

$$\frac{\partial S(p, \lambda)}{\partial s_i} = \frac{1}{1+z} [\theta_i^0 \mathcal{B}_i(p, \lambda) + X_1 \theta_i^1 \mathcal{B}_i(p, \lambda)] 10^{0.4cCL(\lambda)} \times (N_s - i) \lambda_0^{(N_s-i-1)} \quad (\text{A.8})$$

A.2 Modèles d'erreur

Dérivées de l'équation du modèle de variance SALT3, défini 5.31 et en utilisant la définition de la surface g équation 5.42.

Les dérivées du modèle d'erreur par rapport au paramètre du modèle sont les suivantes. On considère M , une donnée spectroscopique (cf. équation 5.28) ou photométrique (cf. équation 5.34), soit :

$$g(p, \lambda) = [\gamma^i \mathcal{G}^i(p, \lambda) * M(p, \lambda)]^2 \quad (\text{A.9})$$

Soit β , les différents paramètres du modèle sauf t_{max} . On peut dériver le modèle d'erreur :

$$\frac{\partial g(p, \lambda)}{\partial \beta} = 2 M(p, \lambda) \frac{\partial M(p, \lambda)}{\partial \beta} [\gamma^i \mathcal{G}^i(p, \lambda)]^2 \quad (\text{A.10})$$

Pour $t_{max}^{B^*}$, il faut ajouter le terme correspondant à la base de spline :

$$\frac{\partial g(p, \lambda)}{\partial t_{max}^{B^*}} = 2 M(p, \lambda) \frac{\partial M(p, \lambda)}{\partial t_{max}^{B^*}} [\gamma^i \mathcal{G}^i(p, \lambda)]^2 + S(p, \lambda)^2 2 \gamma^i \mathcal{G}^i(p, \lambda) \mathcal{G}^i(p, \lambda) \frac{-1}{1+z} \quad (\text{A.11})$$

$$\frac{\partial g(p, \lambda)}{\partial \gamma^i} = 2 \gamma^i [\mathcal{G}^i(p, \lambda) M(p, \lambda)]^2 \quad (\text{A.12})$$

Pour les dérivées de la dispersion en couleur défini équation 5.43, on a :

$$\frac{\partial \sigma^\kappa(\lambda_r^*)}{\partial \sigma_i^\kappa} = i \lambda_r^{*i-1} \quad (\text{A.13})$$

Table des figures

1.1	Mesure des anisotropies de température du fond diffus cosmologique par le satellite PLANCK [Planck Collaboration et al. (2018)]. Les zones bleues correspondent aux régions "froides" et les rouges aux régions, "chaudes". La variation maximale entre deux régions de cette carte est de $600 \mu K$, autour de : $2.7255 K$ [Fixsen (2009)].	16
1.2	Cartographie de plus de 200 000 galaxies réalisée par l'observatoire anglo-australien (AAO) sur une période de 5 ans dont la magnitude limite est de 19.5. On peut voir que deux régions "locales" de l'Univers, entourées en rouge, les densités de galaxies y sont différentes. Tandis que si l'on considère de plus grandes régions, en bleu, la densité de galaxies tend à être similaire.	16
1.3	Diagramme de Hubble [Hubble (1929)], vitesse d'éloignement de 32 galaxies en fonction de leur distance. Première observation de l'expansion de l'Univers.	17
1.4	Distance angulaire rapport de la taille d'un objet D et de son angle apparent $\Delta\theta$.	20
1.5	Distance de luminosité d_L d'une source de luminosité intrinsèque L . L'observateur reçoit un flux surfacique.	21
1.6	Comparaison des différentes distances (calculées pour un même jeu de paramètres cosmologiques) en fonction du <i>redshift</i> . À bas <i>redshift</i> , elles sont asymptotiquement égales.	22
1.7	Courbe de rotation de la galaxie M33 (points bleus). En rouge la somme des différentes contribution du modèle, en tiret-point la contribution du halo de matière noire, en tiret court la contribution du disque stellaire et en tiret long la contribution du gaz interstellaire [Corbelli & Salucci (1999)].	27
1.8	Graphique des limites supérieures de sections efficaces d'interaction en fonction de la masse des particules de matière noire, établies par les différentes expériences.	27
1.9	Diagramme de Hubble avec les galaxies observées par Hubble [Hubble (1929)], jusqu'au <i>redshift</i> de 0.015. Les courbes bleues représentent la prédiction théorique de modèles cosmologiques contenant seulement de la matière et des particules relativistes.	28
1.10	Diagramme de Hubble étendu à partir de donnée de JLA. Les courbes théoriques d'Univers composées seulement de matière ne décrivent pas les observations. . . .	28
1.11	Diagramme de Hubble étendu grâce aux données de JLA avec les prédictions de modèles d'Univers composés de matière, courbes bleues et de modèle d'Univers plats composés de matière et d'énergie noire (sous forme de constante cosmologique), en rouge. Le modèle décrivant le mieux les données est composé à environ 70% d'énergie noire et 30 % de matière (baryonique et noire).	29

1.12	Densité d'énergie en fonction du facteur d'échelle pour les différents fluides composant l'Univers, en vert pour l'énergie noire, en bleu pour la matière et en rouge pour la radiation. Trois phases se succèdent, délimitées par les lignes verticales, une première pour laquelle la densité de particules relativistes domine, suivie d'une phase pour laquelle la matière non-relativiste domine. Aujourd'hui la densité d'énergie noire domine.	31
1.13	Spectre du CMB, les points sont les données mesurées par des expériences spatiales et au sol. La courbe représente le calcul théorique d'un spectre de corps noir à $T \sim 2.7255$ K.	33
1.14	Diagramme de Hubble [Perlmutter et al. (1999)] de 42 SNe jusqu'à un <i>redshift</i> de 0.8. On voit sur le diagramme de résidus que le modèle privilégié est composé à 72 % d'énergie noire sous forme de constante cosmologique et 28 % de matière.	35
1.15	Contours de confiance dans le plan $(\Omega_\Lambda, \Omega_m)$, autour de paramètres qui minimisent les résidus de l'ajustement de la figure 1.14.	35
1.16	Fonction du spectre de puissance des anisotropies de température du CMB (points rouges), modèle Λ CDM ajusté (courbe bleue).	36
1.17	Coupe à différents <i>redshift</i> de l'évolution du profil de masse de chaque espèce en fonction du rayon comobile, extrait de Eisenstein et al. (2007)	38
1.18	Fonction de corrélation dans l'espace des <i>redshifts</i> aux grandes échelles du relevé Large Red Galaxies (LRG) du SDSS. Différents modèles cosmologiques possédant différentes valeurs d'équation d'état de l'énergie noire sont représentés en couleur. Le modèle en magenta est un modèle purement CDM, il manque le pic BAO.	39
1.19	Diagramme de Hubble (équation 1.44) réalisé avec différents traceurs de matière observés à différents <i>redshifts</i> . Figure extraite de Alam et al. (2021)	40
1.20	Contours de confiance dans le plan (w, Ω_m) extrait de Betoule et al. (2014). La zone en bleu correspond aux contraintes imposées par les SNe, en vert celles produites par le CMB et en gris leur combinaison. Figure extraite de [Betoule et al. (2014)].	40
1.21	Contours de confiance dans le plan (w_a, w_0) extrait de Planck Collaboration et al. (2020), on voit que la contrainte la plus précise sur la mesure de (w_0, w_a) est celle imposée par la combinaison CMB + BAO + SNe	40
1.22	Résultats cosmologiques réalisé par Brout et al. (2022) de la combinaison des sondes de CMB, BAO et SNe, pour deux cosmologies : $FlatwCDM$ et $Flatw_0w_aCDM$. 41	
1.24	Illustration de la mesure de distance par parallaxe et loi de Leavitt.	42
1.23	Diagramme de Hubble dont le point zéro est déterminé par une échelle de distance. La distance des Supernovæ est calibrée sur les céphéides. Figure extraite de [Riess et al. (2021)].	43
1.25	Tension sur la valeur de H_0 mesurée par différentes sondes cosmologiques. Les colonnes violette et orange, correspondent respectivement aux valeurs mesurées par Planck et le programme SHOES [Riess et al. (2021)]. Les mesures de l'Univers primordial, se concentre autour de la valeur de Planck dans la colonne violette. Celles inférées par des sondes de l'Univers proche, autour de $74 \text{ km.s}^{-1}.\text{Mpc}^{-1}$, se concentrent dans la colonne orange, on peut notamment citer les SNe Ia mais aussi les mesures de lentillage gravitationnel fort. Figure adaptée de Perivolaropoulos & Skara (2021).	47
2.1	Deux observations de Supernovæ à quatre cent ans d'intervalle.	51
2.2	Spectres des quatre différents types de supernovæ (de haut en bas) SN Ia, SN Ib, SN Ic et SN II, avec les raies des différents éléments chimiques caractéristiques.	53

2.3	Courbes de lumière typiques des grands types de supernovæ. En bleu, la courbe de lumière d'une SN Ia, la luminosité maximale est jusqu'à 10 fois supérieure à celle des autres types de SN. Figure extraite de Baumont (2010).	54
2.4	Proportions de SNe dans un volume fini [Li et al. (2011)].	54
2.5	Taux d'explosion de SN Ia obtenu en combinant les données de différents relevés de SNe : SDSS, SNLS, SUBARU, Palomar & CANDELS. Les distances sont calculées en supposant Λ CDM plat avec $H_0 = 70 \text{ km.s}^{-1}/\text{Mpc}$ et $(\Omega_\Lambda, \Omega_m) = (0.75, 0.25)$	55
2.6	Suivi spectroscopique de SN2011fe et photométrique de SN2008hv. La proximité des objets permet un échantillonnage dense d'observations.	59
2.7	Spectre typique d'une supernova au maximum de luminosité en bande B. Figure extraite de Chotard (2011).	60
2.8	Schématisation de profils P-Cygni, figure extraite de Roth & Kasen (2018).	60
2.9	Schéma de deux paramétrisations de l'évolution temporelle de courbes de lumière en bande B avec le paramètre Δm_{15} en rouge et le stretch, s en bleu. Le modèle en noir peut soit être paramétrisé selon son écart en magnitude entre sa valeur à la phase 0 et +15 jours, soit défini selon le <i>stretch</i> , dilatation temporelle des phases.	62
2.10	Loi de standardisation de 740 SNe Ia extraite de la Joint Light Curve Analysis (JLA) Betoule et al. (2014).	64
2.11	Résidus au diagramme de Hubble de l'analyse JLA [Betoule et al. (2014)] en fonction du logarithme de la masse stellaire de la galaxie hôte. On constate un écart de 0.06 mag entre les deux populations de SNe défini par $M_{\text{stellar}} < 10^{10} M_\odot$ et $M_{\text{stellar}} \geq M_\odot$	65
2.12	Proportions des sous-classes de SNe Ia dans un volume fini [Li et al. (2011)].	66
2.13	Ce graphique, extrait de Taubenberger (2017), montre la diversité des SN Ia. Il représente l'évolution du paramètre de forme Δm_{15} en fonction de la magnitude absolue de représentants des différentes sous-classes des SNe Ia.	67
3.1	Schéma de la collecte et du transfert de charge dans un CCD.	72
3.2	Efficacité quantique de deux caméras, Mégacam utilisée pour les observations du relevé SNLS et HSC caméra du télescope Subaru (cf. Figure 3.13b). Les CCD de cette dernière sont optimisés pour observer dans le rouge.	73
3.3	plan focal de la caméra CFHT12K.	74
3.4	plan focal de la caméra DECam.	74
3.5	plan focal de la caméra Hyper Suprime Cam (HSC)	74
3.6	plan focal de la caméra de l'observatoire Vera Rubin.	74
3.7	À gauche, observation photométrique de la galaxie 19843 du suivi 3D-HST [Brammer et al. (2012)] et à droite l'image dispersée de cette même galaxie.	75
3.8	Trois techniques de spectroscopie à intégrale de champ, figure adaptée de Westmoquette et al. (2009)	76
3.9	Principe de fonctionnement de la spectroscopie à intégrale de champ réalisée par l'instrument SNIFS. Le plan focal est divisé en 225 spaxels. Pour chaque spaxel, deux spectres sont obtenus. De chaque champ, un spectre est obtenu permettant de constituer un cube de données à deux dimensions spatiales et une spectrale pour chaque pose. Les spectres des SNe en sont extraits. Le suivi réalisé permet de synthétiser des courbes de lumière pour ces objets. Figure extraite de Chotard (2011).	77
3.10	Transmission des filtres $UBVRI$ et $ugriz$, avec un spectre d'une SN observé à bas z proche du maximum de luminosité en bande B.	78

3.11	Observation photométrique du champs COSMOS, par l'instrument HyperSupprimeCam (HSC, de plan focale visible figure 3.5) du télescope Subaru (cf. Figure 3.13b)	80
3.12	Composants de l'ajustement du modèle de PSF pour une image dans la bande g_M proche du maximum de la SN Ia 03D4ag ($z = 0.285$) du relevé SNLS. Figure extraite de Fourmanoit (2010).	82
3.14	Diagramme de Hubble extrait de l'analyse Panthéon Scolnic et al. (2018). Combinaison de 1048 SNe de <i>redshift</i> compris entre 0.03 et 2.3, observées par 12 relevés différents.	92
3.15	Proportion d'erreurs systématiques sur l'estimation de l'équation d'état de l'énergie noire [Scolnic et al. (2018)]. La source principale d'erreurs systématiques est la propagation des erreurs de calibration dans le "modèle d'ajustement de courbes de lumière" utilisé nommé SALT2 (dont on verra, le but, le développement, les réussites et les limitations dans la suite du manuscrit).	93
3.16	Le biais de Malmquist affecte les distributions de stretch et de couleur des SNe à grand <i>redshift</i> de JLA.	95
4.1	Spectres de quatre SNe émis à la même phase et à des <i>redshift</i> de 0.1, 0.4, 0.7, et 1. Quatre bandes <i>griz</i> sont définies dans le référentiel de l'observateur. L'intégrale d'un spectre dans cette bande correspond à la magnitude observée. La partie surlignée en bleu de chaque spectre correspond au filtre B^* défini dans le référentiel de chaque SN. L'intégrale de cette partie de chaque spectre correspond à la quantité nécessaire pour inférer des distances. Pour un spectre émis à $z = 0.1$, cette quantité est observée en bande g . Pour un spectre émis à $z = 1$, elle est observée dans les bandes i, z	102
4.2	Surface spectrale et coupes à phase fixée. Figure extraite de Hsiao et al. (2007) .	104
4.3	Ingrédients du modèles SiFTO évalué pour trois points de photométrie de la SN 06D2ag. En arrière plan, les transmissions des trois filtres (à ce <i>redshift</i> $z \sim 0.31$, l'information contenue dans bande z est hors des limites du domaine de définition du modèle) de MegaCam, g , r et i . Les points de photométrie sont représentés avec les barres d'erreur en noir, en tirets, le spectre, extrait du modèle de référence ϕ_0 . Les cercles creux noirs sont le flux intégré de ce spectre dans les différentes bandes. En rouge, la SED ajustée en couleur, grâce à la fonction de splines S (en vert) et dont l'intégrale à travers chaque bande donne les points de photométrie (cercles bleus).	108
4.4	Corrélation entre les paramètres de SNEMO2 et SALT2, extraits sur le même lot de données. Les points rouges correspondent aux SNe définies comme particulières. Figure extraite de Saunders et al. (2018).	110
4.5	Coefficients de corrélation entre les paramètres de SUGAR et SALT2. Figure extraite de Léget et al. (2020).	112
4.6	Spectres de trois SNe Ia appartenant à trois sous-classes différentes : normale à gauche, 91T-like au milieu et 91-bg-like à droite. Les courbes bleues représentent le modèle issu des deux premières étape sans paramétrisation non-linéaire de la variabilité intrinsèque. Les courbes rouges correspondent à un modèle de carte iso possédant trois paramètres et ajusté par processus gaussiens. Figure adapté de Boone (2021)	113
4.7	Matrices de covariance des résidus des spectres de SNe Ia pour deux modèles. À gauche le modèle estimé à la première étape corrigé des effets extrinsèques. À droite le modèle de carte iso avec trois paramètres. Figure extraite de Boone (2021).113	

4.8	Coefficients de corrélation entre les trois paramètres du modèles (colonnes) en fonction de caractéristiques spectrales et des paramètres de SALT2, SNEMO et SUGAR (lignes). Figure extraite de Boone (2021)	114
4.9	Histogramme des <i>redshifts</i> de SNe composant les lots d’entraînements de JLA et de l’ajout de K21 (figure extraite de Kenworthy et al. (2021)).	115
5.1	Surface M_0 du modèle SALT2.4, en fonction de la phase et de la longueur d’onde.	121
5.2	Surface M_1 du modèle SALT2.4, en fonction de la phase et de la longueur d’onde.	122
5.3	Loi de couleur du modèle SALT2.4 en fonction de la longueur d’onde.	122
5.4	Visualisation des différentes composantes spectrales. La colonne de gauche l’effet de la prise en compte successive des contributions de M_0 , M_1 et CL pour modéliser un spectre (points noir). Sur la colonne de droite, on peut voir la valeur des différentes contributions indépendamment.	137
5.5	Représentation de la modélisation de courbes de lumières. Création de deux courbes de lumière générées à travers deux filtres g (bleu) et i (rouge) de l’instrument du futur télescope du l’observatoire Vera Rubin.	138
5.6	Dispersion en couleur modélisé par SALT2 (en bleu) et SALT3 (en rouge). La paramétrisation du terme de dispersion en couleur, dépend du lot d’entraînement. Le modèle SALT2 paramétrise cette variance additionnel par un polynôme de degré 2, le modèle SALT3 quand à lui, par un polynôme de degré 3. Figure extraite de Kenworthy et al. (2021)	139
5.7	Bases de B-splines d’ordre 1 (graphique du haut), d’ordre 2 (graphique du milieu) et d’ordre 3 (graphique du bas), avec 11 (courbes pleines) et 5 (courbes en tirets) nœuds. Plus l’ordre augmente, plus il y a de recouvrement entre les splines. Ici les graphique ont été réalisés avec des paramètres tous égaux à un. Plus le nombre de nœud est grand, plus complexes sont les formes pouvant être décrites.	140
5.8	Nombre de paramètres du modèle en fonction du nombre de SN. Le nombre de spectres par SN varie entre 1 et 5. Les deux étoiles correspondent aux nombres de SNe et de spectres contenus dans les lots d’entraînement de SALT2.4 (JLA) et SALT3 (K21).	141
5.9	Temps d’évaluation moyen, du modèle d’une courbe de lumière (panneau du haut), des dérivées du modèle d’une courbe de lumière (2ème panneau), du modèle d’un spectre (3ème panneau) et des dérivées du modèle d’un spectre (dernier panneau) en fonction du nombre de SNe dans le lot d’entraînement.	142
6.1	Impact du coefficient de régularisation, μ_{reg} , sur l’ajustement d’une courbe de lumière par des splines. Les points noirs correspondent à des courbes de lumières simulées avec différents échantillonnages, 8 points à gauche et 50 points à droites. Pour les modèles rouges la régularisation pénalise les paramètres non contraints. Pour les modèles bleus, la régularisation pénalise aussi la pente entre les paramètres consécutifs en phase.	149
7.1	Schéma du module de simulation. Il est composé de trois étapes. Lors de la première : on crée les dates et les caractéristiques des observations et les paramètres propres à chaque SN, en supposant une cosmologie pour les simulations idéales et pour les lots réalistes, les observations sont recopiées et les paramètres d’un ajustement de SALT2.4 sont utilisés. La seconde étape est la modélisation du flux, en utilisant le modèle SALT2.4. La dernière étape ajoute du bruit aux spectres et courbes de lumière.	160

7.2	Distribution des phases spectrales du lot d'entraînement de K21. Les cinq courbes représentent différentes fonctions de densité ajustées sur l'histogramme. La courbe verte correspondant à un ajustement minimal, représente une distribution de Gumbel dont les paramètres sont $\mu = 2.5$ et $\beta = 12.8$	162
7.3	Distribution des maxima de luminosité de chaque relevé composant JLA (haut) et K21 (bas).	164
7.4	Distribution du nombre de spectres par supernova composant chaque relevé du lot d'entraînement de JLA.	165
7.5	Distribution du nombre de spectres par supernova composant chaque relevé du lot d'entraînement de K21.	166
7.6	Nombre de courbes de lumières par SNe en fonction des relevés composant JLA.	166
7.7	Nombre de courbes de lumières par SNe en fonction des relevés composant K21.	167
7.8	Nombre de points par courbe de lumière en fonction des relevés composant JLA.	168
7.9	Nombre de points par courbe de lumière en fonction des relevés composant K21.	169
7.10	Diagramme de couverture des données photométriques et spectroscopiques de JLA.	170
7.11	Diagramme de couverture des données photométriques et spectroscopiques de K12.	170
7.12	Exemple de coupure en longueur d'onde pour un modèle défini entre 200 et 900 nm.	171
7.13	Comparaison des données (points) de la SN 1994s à $z = 0.015$, avec les simulations (triangle), pour les données photométriques à gauche et les spectres à droite.	173
7.14	Comparaison des données (points) de la SN05D4av observé par le relevé SNLS à $z = 0.509$, avec les simulations (triangle).	174
8.1	État initial du modèle de spectre à gauche (courbe bleu) et à droite pour les courbes de lumière des 5 bandes. Les points représentent les données simulées. La SN est simulée à $z = 0.59$, elle n'a donc que quatre courbes de lumière.	177
8.2	Distribution des résidus dé-pesés de l'erreur à l'état initial de l'ajustement. La figure de gauche représente les résidus de l'ensemble des données, celle du milieu des données photométriques et celle de droite des données spectroscopiques. Les courbes représente la distribution d'une loi gaussienne dont les paramètres sont extraits des histogrammes et affichées dans la légende.	178
8.3	Évolution des différentes contributions du χ^2 et des contraintes pour chacune des 10 itérations.	180
8.4	Reconstruction des paramètres de couleur et X_1 pour 500 simulations LSST.	181
8.5	Reconstruction des paramètres η pour 500 simulations LSST.	181
8.6	Reconstruction de la dispersion en couleur pour une simulation LSST de 500 SNe. Le panneau du haut montre la distribution des λ_r^* . Le panneau du milieu représente la reconstruction des paramètres κ et de la dispersion en couleur. Les grandeurs noires sont les valeurs utilisées pour la simulation et les rouges les valeurs reconstruites. Les paramètres de chaque polynômes sont indiqué dans la légende (de gauche à droite, le paramètres d'ordre le plus grand au plus petit). Le panneau inférieur montre les résidus. Les amplitudes des courbes de lumière dans l'UV, $\lambda_{red}^* < -1$ sont bien reconstruites car moins bien mesurées.	182
8.7	Reconstruction des paramètres du modèle et comparaison avec les paramètres de SALT2.4. Le panneau du haut montre la surface reconstruite, celui du milieu la surface utilisée pour générer les simulations, celui du bas leur rapport. La loi de couleur reconstruite est affichée en bleu, celle de SALT2.4 en rouge et le ratio des deux est affiché sur le panneau inférieur.	182
8.8	Ajustement de notre SN avec le modèle obtenu à l'issue du premier ajustement.	183

8.9	Distribution des <i>pulls</i> à l'issue du premier ajustement, à gauche de toutes les données, au milieu des données photométriques et enfin à droite des données spectroscopiques. Le trait plein correspond à l'histogramme des <i>pulls</i> et la courbe en tiret correspond à une fonction de densité gaussienne dont les paramètres sont indiqués dans la légende. Ils sont centrés mais non réduits, comme attendu puisque seul le modèle et le <i>color scatter</i> sont ajustés.	184
8.10	Ajustement du modèle et du modèle d'erreur obtenus à l'issue du troisième ajustement sur notre SN.	185
8.11	Distribution des <i>pulls</i> à l'issue du troisième ajustement. Ils sont maintenant centrés et réduits.	186
8.12	Reconstruction du modèle de la surface du modèle d'erreur.	186
8.13	<i>Pull</i> de l'ajustement du modèle sur les courbes de lumière d'une simulation de 1000 SNe du relevé LSST par bande d'observation.	187
8.14	<i>Pull</i> par intervalles de phases de l'ajustement du modèle sur les spectres d'une simulations de 1000 SNe de LSST.	187
8.15	Temps de calcul par itération lorsque tous les paramètres sont ajustés simultanément pour 5 lots d'entraînement, respectivement de 50, 100, 250, 500, 1000 et 5000 SNe.	188
8.16	Temps de calcul des principales composantes du minimiseur pour différentes simulations. De gauche à droite pour 250, 500, JLA, 1000, K21 et 5000. Les principales composantes sont le temps passé dans l'inversion de la matrice hessienne, dans la factorisation de <i>cholesky</i> ; le temps passé à sa construction, la taille des matrices rend les produits matriciels lents et le temps passé à évaluer les modèles et leurs dérivées.	191
8.17	<i>Pull</i> de l'ajustement du modèle sur les courbe de lumière observées par SNLS de la simulation JLA.	192
8.18	<i>Pull</i> de l'ajustement du modèle sur les courbe de lumière observées par les relevés STANDARD de la simulation JLA.	193
8.19	<i>Pull</i> par intervalle de phase de l'ajustement du modèle sur les spectres de la simulation JLA.	194
8.20	Reconstruction des paramètres de calibration η par relevé de JLA sur la panneau de gauche, avec les incertitudes à 1σ en rouge. Sur le panneau du milieu est représenté la somme des paramètres η par relevé. Enfin le panneau de droite montre les paramètres par instrument rangé en couleur.	195
8.21	Ajustement du <i>color scatter</i> et X_0 sur une simulation JLA, lorsque tous les autres paramètres sont fixés à la vérité.	196
8.22	Simulation de la SN 1995bd observée à $z \sim 0.015$ par les relevés historiques. Reconstruction du modèle et du modèle d'erreur.	197
8.23	Simulation de la SN 05D2bv observée à $z \sim 0.63$ par SNLS. Reconstruction du modèle et du modèle d'erreur.	198
8.24	Simulation de la SN ASASSN-15mg observée à $z \sim 0.0428$ par Pan-Starr. Reconstruction du modèle et du modèle d'erreur.	200
9.1	Reconstruction des biais sur les paramètres des SNe pour une simulation LSST de 500 supernovæ.	204
9.2	Reconstruction des biais sur les paramètres des SNe pour une simulation LSST de 1000 supernovæ.	205
9.3	Reconstruction des biais sur les paramètres des SNe pour les 50 simulations LSST de 500 supernovæ.	206

9.4	Reconstruction des biais sur les paramètres des SNe pour 40 simulations LSST de 500 supernovæ, pour lequel le paramètre de régularisation à été divisé par 10 lors du troisième ajustement.	207
9.5	Reconstruction des biais sur les paramètres des SNe pour 50 simulations LSST de 1000 supernovæ.	208
9.6	Reconstruction des biais sur les paramètres des SNe pour 50 simulations LSST de 5000 supernovæ.	209
9.7	Comparaison de incertitudes sur les paramètres d'intérêt ($X_0 : 0$ à 1000 , $X_1 : 1$ à 2000 et $c : 2000$ à 3000) extraites des matrices de Fisher obtenues pour les 50 ajustements.	211
9.8	Test de couverture pour les couleurs X_1 , pour l'ajustement du modèle sur les simulations LSST de 1000 SNe.	212
9.9	Test de couverture pour les paramètres de forme c , pour l'ajustement du modèle sur les simulations LSST de 1000 SNe.	213
9.10	Test de couverture pour la magnitude au maximum de luminosité en bande B^* , pour l'ajustement du modèle sur les simulations LSST de 1000 SNe.	214
9.11	Test de couverture pour les module de distance μ , pour l'ajustement du modèle sur les simulations LSST de 1000 SNe.	215
9.12	Test de couverture pour la magnitude au maximum de luminosité en bande B^* , pour l'ajustement d'un modèle sur les simulations LSST de 1000 SNe, calculé avec une matrice de Fisher pour laquelle le terme constant de la dispersion en couleur est dé-biaisé.	216
9.13	Test de couverture pour les module de distance μ , pour l'ajustement d'un modèle sur les simulations LSST de 1000 SNe, calculé avec une matrice de Fisher pour laquelle le terme constant de la dispersion en couleur est dé-biaisé.	217
9.14	Biais sur les paramètres d'intérêts issu de l'ajustement du modèle et de l' <i>error snake</i> sur JLA. Pour chaque paramètre, m_{B^*} , X_1 , c sont les erreur extraites directement de la matrice de Fisher.	219
9.15	Biais sur les paramètres d'intérêts issu de l'ajustement du modèle et de l' <i>error snake</i> sur JLA. À gauche pour l'ensemble des simulations pour lesquelles le prior de calibration est tenu au pour mille, à droite quand il est donné par la matrice de calibration de JLA.	220
9.16	Test de couverture pour la magnitude au maximum de luminosité en bande B^* , pour l'ajustement d'un modèle propageant les incertitudes de calibration de JLA sur le lot d'entraînement de JLA. En haut la matrice $V^F(\mu)$ définie équation 9.6), en dessous la matrice de covariance empirique $V^E(\mu)$ (cf. équation 9.8). Sur le panneau supérieur du milieu, on peut voir les déviations standard sur les magnitudes, racine carré de la diagonale des matrices de covariances en fonction du <i>redshit</i> . Les déviations standard de Fisher issu de la hessien de l'ajustement sont représentées par des triangles, celles empiriques sont représentées par des triangles noire. Le code couleur code les différents relevés de supernovæ composant JLA. En dessous, sont présentés les rapports de ces estimations, dont la valeur moyenne est 1. À droite, sont affichées les histogrammes correspondant.	221
9.17	Test de couverture pour les paramètres de forme X_1 , pour l'ajustement d'un modèle propageant les incertitudes de calibration de JLA sur le lot d'entraînement de JLA. Le ratio des déviations standards entre les triangles colorés (issu de la matrice de Fisher) et les déviations standards empiriques, en noirs a une valeur moyenne est 1.	222

9.18 Test de couverture pour les paramètres de couleur c , pour l'ajustement d'un modèle propageant les incertitudes de calibration de JLA sur le lot d'entraînement de JLA. Le ratio des déviations standards entre les triangles colorés (issu de la matrice de Fisher) et les déviations standards empiriques, en noirs a une valeur moyenne est 1.223

9.19 Test de couverture pour les modules de distance μ , pour l'ajustement d'un modèle propageant les incertitudes de calibration de JLA sur le lot d'entraînement de JLA. Le rapport des déviations standards entre les triangles colorés (issu de la matrice de Fisher) et les déviations standards empiriques, en noirs a une valeur moyenne est 1, à 13% près. 224

Liste des tableaux

3.1	Données photométriques et spectroscopiques des SNe Ia (de <i>redshift</i> connu) observées et suivies par CFA, CSP, SNLS et SDSS, à notre disposition. Un travail de récolte, de compréhension et de reformatage de ces données a été réalisé. . . .	90
5.1	Contraintes explicites sur les paramètres du modèle.	126
5.2	Nombre de paramètres du modèle en fonction du nombre de SNe : N_{SN} , du nombre de courbes de lumière : N_{lcs} , du nombre de spectres : $N_{spectres}$ et du nombre de bandes : N_{bandes}	133
7.1	Code des noms des différents relevés utilisés dans les lots JLA et K21.	163
8.1	Nombre d'itérations, durée d'une itération et temps total passé dans chacun des trois ajustements du processus d'entraînement.	189
8.2	Incertitude de calibration sur les instruments des différents relevés utilisés dans l'analyse JLA.	194

Bibliographie

- 2004, Parallaxe annuelle, https://commons.wikimedia.org/wiki/File:Parallaxe_annuelle.png
- Aguilar-Arevalo, A., Amidei, D., Bertou, X., et al. 2015, arXiv e-prints, arXiv:1510.02126
- Alam, S., Ata, M., Bailey, S., et al. 2017, *Monthly Notices of the Royal Astronomical Society*, 470, 2617
- Alam, S., Aubert, M., Avila, S., et al. 2021, *Phys. Rev. D*, 103, 083533
- Aldering, G., Adam, G., Antilogus, P., et al. 2002, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4836, Survey and Other Telescope Technologies and Discoveries, ed. J. A. Tyson & S. Wolff, 61–72
- Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, *Annual Review of Astronomy and Astrophysics*, 49, 409
- Alsabti, A. W. & Murdin, P. 2017, Handbook of Supernovae
- Aprile, E., Aalbers, J., Agostini, F., et al. 2017, *European Physical Journal C*, 77, 881
- Arnett, W. D. 1982, *Astrophysical Journal*, 253, 785
- Astier, P., El Hage, P., Guy, J., et al. 2013, *Astronomy and Astrophysics*, 557, A55
- Astier, P., Guy, J., Regnault, N., et al. 2006, *Astronomy and Astrophysics*, 447, 31
- Balland, C., Baumont, S., Basa, S., et al. 2009, *Astronomy and Astrophysics*, 507, 85
- Barone-Nugent, R. L., Lidman, C., Wyithe, J. S. B., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 425, 1007
- Baumont, S. 2010, Analyse des spectres VLT pour l'expérience SNLS Qualification de transients cosmologiques, Phd
- Bautista, J. E., Paviot, R., Vargas Magaña, M., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 500, 736
- Bessell, M. S. 1990, *Publications of the Astronomical Society of the Pacific*, 102, 1181
- Betoule, M., Kessler, R., Guy, J., et al. 2014, *Astronomy and Astrophysics*, 568, A22
- Betoule, M., Murriner, J., Regnault, N., et al. 2013, *Astronomy and Astrophysics*, 552, A124
- Blondin, S., Matheson, T., Kirshner, R. P., et al. 2012, *Astrophysical Journal*, 143, 126
- Bloom, J. S., Kasen, D., Shen, K. J., et al. 2012, *Astrophysical Journal Letters*, 744, L17
- Bohlin, R. C. 2014, *Astrophysical Journal*, 147, 127
- Boone, K. 2021, *Astrophysical Journal*, 162, 275
- Boone, K., Aldering, G., Antilogus, P., et al. 2021, *ApJ*, 912, 71
- Boulade, O., Charlot, X., Abbon, P., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, ed. M. Iye & A. F. M. Moorwood, 72–81
- Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, *ApJS*, 200, 13

-
- Branch, D. ; Garnavich, P., Matheson, T. ; Baron, E., Thomas, R. C., et al. 2003, *The Astronomical Journal Supplement*, 126, 1489
- Bronder, T. J., Hook, I. M., Astier, P., et al. 2008, *Astronomy and Astrophysics*, 477, 717
- Brout, D., Scolnic, D., Popovic, B., et al. 2022, *arXiv e-prints*, arXiv:2202.04077
- Caldwell, R., Dave, R., & Steinhardt, P. J. 1998, *Physical Review Letters*, 80, 1582
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *Astrophysical Journal*, 345, 245
- Chabanier, S., Palanque-Delabrouille, N., Yèche, C., et al. 2019, *Journal of Cosmology and Astroparticle Physics*, 2019, 017
- Chandrasekhar, S. 1931, *Astrophysical Journal*, 74, 81
- Childress, M., Aldering, G., Antilogus, P., et al. 2013, *Astrophysical Journal*, 770, 108
- Chon, G., Böhringer, H., Collins, C. A., & Krause, M. 2014, *Astronomy and Astrophysics*, 567, A144
- Chotard, N. 2011, Etude de la variabilité des Supernovae de type Ia observées par la collaboration Nearby Supernova Factory, Phd, Université Claude Bernard
- Chotard, N., Gangler, E., Aldering, G., et al. 2011, *Astronomy and Astrophysics*, 529, L4
- Colgate, S. A. & McKee, C. 1969, *Astrophysical Journal*, 157, 623
- Colless, M., Dalton, G., Maddox, S., et al. 2001, *Monthly Notices of the Royal Astronomical Society*, 328, 1039
- Conley, A., Guy, J., Sullivan, M., et al. 2011, *Astrophysical Journal Supplement Series*, 192, 1
- Conley, A., Sullivan, M., Hsiao, E. Y., et al. 2008, *Astrophysical Journal*, 681, 482
- Corbelli, E. & Salucci, P. 1999, *Monthly Notices of the Royal Astronomical Society* [[arXiv]astro-ph/9909252]
- de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, *The Messenger*, 175, 3
- de Jong, R. S., Barden, S., Bellido-Tirado, O., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, ed. S. K. Ramsay, I. S. McLean, & H. Takami, 91470M
- DePoy, D. L., Abbott, T., Annis, J., et al. 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II, ed. I. S. McLean & M. M. Casali, 70140E
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, *arXiv e-prints*, arXiv:1611.00036
- Dhawan, S., Goobar, A., Smith, M., et al. 2022, *Monthly Notices of the Royal Astronomical Society*, 510, 2228
- Di Valentino, E., Mena, O., Pan, S., et al. 2021, *Classical and Quantum Gravity*, 38, 153001
- Einstein, A. 1915, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, Seite 844-847., 844
- Einstein, A. 1917, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, Seite 142-152., 142
-

- Eisenstein, D. J., Seo, H., & White, M. 2007, *The Astrophysical Journal*, 664 [[arXiv]0604361]
- Elias, J. H., Matthews, K., Neugebauer, G., & Persson, S. E. 1985, *Astrophysical Journal*, 296, 379
- Ellis, R. S., Sullivan, M., Nugent, P. E., et al. 2008, *Astrophysical Journal*, 674, 51
- Fabricant, D., Cheimets, P., Caldwell, N., & Geary, J. 1998, *Publications of the Astronomical Society of the Pacific*, 110, 79
- Fakhouri, H. K., Boone, K., Aldering, G., et al. 2015, *Astrophysical Journal*, 815, 58
- Filippenko, A. V., Richmond, M. W., Branch, D., et al. 1992a, *Astrophysical Journal*, 104, 1543
- Filippenko, A. V., Richmond, M. W., Matheson, T., et al. 1992b, *Astrophysical Journal Letters*, 384, L15
- Fisher, R. A. 1992, *Statistical Methods for Research Workers*, ed. S. Kotz & N. L. Johnson (New York, NY : Springer New York), 66–70
- Fixsen, D. 2009, *Astrophysical Journal*, 707, 916
- Folatelli, G., Morrell, N., Phillips, M. M., et al. 2013, *Astrophysical Journal*, 773, 53
- Foley, R. J., Chornock, R., Filippenko, A. V., et al. 2009, *Astrophysical Journal*, 138, 376
- Foley, R. J., Filippenko, A. V., Kessler, R., et al. 2012, *Astrophysical Journal*, 143, 113
- Foley, R. J., Scolnic, D., Rest, A., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 475, 193
- Fourmanoit, N. 2010, *Analyse des 5 ans de données de l'expérience SuperNova Legacy Survey*, Phd, Université Pierre et Marie Curie
- Freedman, W. L., Madore, B. F., Hatt, D., et al. 2019, *Astrophysical Journal*, 882, 34
- Frieman, J. A., Bassett, B., Becker, A., et al. 2008, *Astronomical Journal*, 135, 338
- Frohmaier, C., Sullivan, M., Nugent, P. E., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 486, 2308
- Fukuda, Y., Hayakawa, T., Ichihara, E., et al. 1998, *Phys. Rev. Lett.*, 81, 1562
- Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, *Astrophysical Journal*, 111, 1748
- Graziani, R., Rigault, M., Regnault, N., et al. 2020, *arXiv e-prints*, arXiv:2001.09095
- Greenstein, J. L. & Schmidt, M. 1964, *Astrophysical Journal*, 140, 1
- Gris, P., Regnault, N., Awan, H., et al. 2022, *arXiv e-prints*, arXiv:2205.07651
- Guy, J., Astier, P., Baumont, S., et al. 2007, *Astronomy and Astrophysics*, 466, 11
- Guy, J., Astier, P., Nobili, S., Regnault, N., & Pain, R. 2005, *Astronomy and Astrophysics*, 443, 781
- Guy, J., Sullivan, M., Conley, A., et al. 2010, *Astronomy and Astrophysics*, 523, A7
- Hamuy, M., Folatelli, G., Morrell, N. I., et al. 2006, *The Publications of the Astronomical Society of the Pacific*, 118, 2
- Hamuy, M., Maza, J., Phillips, M. M., et al. 1993a, *Astrophysical Journal*, 106, 2392

-
- Hamuy, M., Phillips, M. M., Maza, J., et al. 1995, *AJ*, 109, 1
- Hamuy, M., Phillips, M. M., Suntzeff, N. B., et al. 1996, *Astrophysical Journal*, 112, 2408
- Hamuy, M., Phillips, M. M., Wells, L. A., & Maza, J. 1993b, *Publications of the Astronomical Society of the Pacific*, 105, 787
- Hazenbergh, F. 2019, Calibration photométrique des supernovae de type Ia pour la caractérisation de l'énergie noire avec l'expérience StarDICE, Phd, Sorbonne Université
- Henderson, C. R., Kempthorne, O., Searle, S. R., & von Krosigk, C. M. 1959, *Biometrics*, 15, 192
- Hicken, M., Challis, P., Kirshner, R. P., et al. 2012, *The Astronomical Journal Supplement*, 200, 15
- Hicken, M., Garnavich, P. M., Prieto, J. L., et al. 2007, *The Astronomical Journal*, 669, L17
- Hicken, M., Wood-Vasey, W. M., Blondin, S., et al. 2009, *The Astronomical Journal*, 700, 1097
- Hogg, D. W. 1999, *arXiv e-prints*, astro
- Hogg, D. W., Eisenstein, D. J., Blanton, M. R., et al. 2005, *Astrophysical Journal*, 624, 54
- Holtzman, J. A., Marriner, J., Kessler, R., et al. 2008, *Astrophysical Journal*, 136, 2306
- Hou, J., Sánchez, A. G., Ross, A. J., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 500, 1201
- Hounsell, R., Scolnic, D., Foley, R. J., et al. 2018, *Astrophysical Journal*, 867, 23
- Howell, D. A., Sullivan, M., Conley, A., & Carlberg, R. 2007, *Astrophysical Journal Letters*, 667, L37
- Howell, D. A., Sullivan, M., Perrett, K., et al. 2005, *Astrophysical Journal*, 634, 1190
- Hsiao, E. Y., Conley, A., Howell, D. A., et al. 2007, *Astrophysical Journal*, 663, 1187
- Hubble, E. 1929, *Proceedings of the National Academy of Science*, 15, 168
- Humason, M. L., Mayall, N. U., & Sandage, A. R. 1956, *Astrophysical Journal*, 61, 97
- Iben, I., J. & Tutukov, A. V. 1984, *Astrophysical Journal Supplement Series*, 54, 335
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *Astrophysical Journal*, 873, 111
- Jha, S., Riess, A. G., & Kirshner, R. P. 2007, *Astrophysical Journal*, 659, 122
- Jönsson, J., Sullivan, M., Hook, I., et al. 2010, *Monthly Notices of the Royal Astronomical Society*, 405, 535
- Joyce, A., Lombriser, L., & Schmidt, F. 2016, *Annual Review of Nuclear and Particle Science*, 66, 95
- Kelly, P. L., Hicken, M., Burke, D. L., Mandel, K. S., & Kirshner, R. P. 2010, *Astrophysical Journal*, 715, 743
- Kenworthy, W. D., Jones, D. O., Dai, M., et al. 2021, *arXiv e-prints*, arXiv:2104.07795
- Kessler, R., Bernstein, J. P., Cinabro, D., et al. 2009, *Publications of the Astronomical Society of the Pacific*, 121, 1028
-

- Kessler, R., Guy, J., Marriner, J., et al. 2013, *Astrophysical Journal*, 764, 48
- Kessler, R. & Scolnic, D. 2017, *ApJ*, 836, 56
- Khokhlov, A., Mueller, E., & Hoefflich, P. 1993, *Astronomy and Astrophysics*, 270, 223
- Krisciunas, K., Contreras, C., Burns, C. R., et al. 2017, *The Astronomical Journal*, 153, 211
- Krisciunas, K., Li, W., Matheson, T., et al. 2011, *The Astronomical Journal*, 142, 74
- Lampeitl, H., Smith, M., Nichol, R. C., et al. 2010, *Astrophysical Journal*, 722, 566
- Lantz, B., Aldering, G., Antilogus, P., et al. 2004, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 5249, *Optical Design and Engineering*, ed. L. Mazuray, P. J. Rogers, & R. Wartmann, 146–155
- Laureijs, R., Hoar, J., Buenadicha, G., et al. 2014, in Astronomical Society of the Pacific Conference Series, Vol. 485, *Astronomical Data Analysis Software and Systems XXIII*, ed. N. Manset & P. Forshay, 495
- Leavitt, H. S. & Pickering, E. C. 1912, *Harvard College Observatory Circular*, 173, 1
- Léget, P. F., Gangler, E., Mondon, F., et al. 2020, *Astronomy and Astrophysics*, 636, A46
- Lemaître, G. 1927, *Annales de la Société Scientifique de Bruxelles*, 49
- Li, W., Filippenko, A. V., Chornock, R., et al. 2003, *Publications of the Astronomical Society of the Pacific*, 115, 453
- Li, W., Leaman, J., Chornock, R., et al. 2011, *Monthly Notices of the Royal Astronomical Society*, 412, 1441
- LSST DESC. 2012, *arXiv e-prints*, arXiv:1211.0310
- Léna, P. 1996, *Astrophysique : Méthodes physiques de l'observation*
- Macaulay, E., Nichol, R. C., Bacon, D., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 486, 2184
- Mandelbaum, R. 2018, *Annual Review of Astronomy and Astrophysics*, 56, 393
- Marriner, J., Bernstein, J. P., Kessler, R., et al. 2011, *ApJ*, 740, 72
- Matheson, T., Kirshner, R. P., Challis, P., et al. 2008, *Astrophysical Journal*, 135, 1598
- Mayet, F., Green, A. M., Battat, J. B. R., et al. 2016, *Physics Reports*, 627, 1
- Miknaitis, G., Pignata, G., Rest, A., et al. 2007, *Astrophysical Journal*, 666, 674
- Miyazaki, S., Komiyama, Y., Nakaya, H., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, *Ground-based and Airborne Instrumentation for Astronomy IV*, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 84460Z
- Mosher, J., Guy, J., Kessler, R., et al. 2014, *Astrophysical Journal*, 793, 16
- Nomoto, K. 1982, *Astrophysical Journal*, 253, 798
- Nugent, P., Kim, A., & Perlmutter, S. 2002, *Publications of the Astronomical Society of the Pacific*, 114, 803
- Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, *Astrophysical Journal*, 819, 129

-
- Oke, J. B. & Sandage, A. 1968, *Astrophysical Journal*, 154, 21
- Penzias, A. A. & Wilson, R. W. 1965, *Astrophysical Journal*, 142, 419
- Pereira, R., Thomas, R. C., Aldering, G., et al. 2013, *Astronomy and Astrophysics*, 554, A27
- Perivolaropoulos, L. & Skara, F. 2021, arXiv e-prints, arXiv:2105.05208
- Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *Astrophysical Journal*, 517, 565
- Phillips, M. M. 1993, *Astrophysical Journal Letters*, 413, L105
- Phillips, M. M., Li, W., Frieman, J. A., et al. 2007, *Publications of the Astronomical Society of the Pacific*, 119, 360
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, ArXiv e-prints [[arXiv]1807.06205]
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, *ArXiv e-prints*, 641, 16
- Pskovskii, I. P. 1977, *Soviet Astronomy*, 21, 675
- Rest, A., Scolnic, D., Foley, R. J., et al. 2014, *ApJ*, 795, 44
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *Astrophysical Journal*, 116, 1009
- Riess, A. G., Kirshner, R. P., Schmidt, B. P., et al. 1999, *Astrophysical Journal*, 117, 707
- Riess, A. G., Press, W. H., & Kirshner, R. P. 1996, *Astrophysical Journal*, 473, 88
- Riess, A. G., Yuan, W., Macri, L. M., et al. 2021, arXiv e-prints, arXiv:2112.04510
- Rigault, M., Brinnel, V., Aldering, G., et al. 2020, *Astronomy and Astrophysics*, 644, A176
- Roman, M., Hardin, D., Betoule, M., et al. 2018, *Astronomy and Astrophysics*, 615, A68
- Röpke, F. K. & Hillebrandt, W. 2004, *Astronomy and Astrophysics*, 420, L1
- Rose, B. M., Baltay, C., Hounsell, R., et al. 2021, arXiv e-prints, arXiv:2111.03081
- Ross, N. P., Hamann, F., Zakamska, N. L., et al. 2015, *Monthly Notices of the Royal Astronomical Society*, 453, 3932
- Roth, N. & Kasen, D. 2018, *Astrophysical Journal*, 855, 54
- Rubin, D., Aldering, G., Antilogus, P., et al. 2022, arXiv e-prints, arXiv:2205.01116
- Rubin, V. C. & Ford, W. K. J. 1970, *The Astrophysical Journal*, 159, 379
- Rust, B. W. 1974, *The Use of Supernovae Light Curves for Testing the Expansion Hypothesis and Other Cosmological Relations.*, PhD thesis, Oak Ridge National Lab., TN.
- Sako, M., Bassett, B., Becker, A. C., et al. 2014, ArXiv e-prints [[arXiv]1401.3317]
- Sauer, D. N., Mazzali, P. A., Blondin, S., et al. 2008, *Monthly Notices of the Royal Astronomical Society*, 391, 1605
- Saunders, C., Aldering, G., Antilogus, P., et al. 2018, *ApJ*, 869, 167
- Schlafly, E. F. & Finkbeiner, D. P. 2011, *Astrophysical Journal*, 737, 103
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *Astrophysical Journal*, 500, 525
- Scolnic, D., Casertano, S., Riess, A., et al. 2015, *Astrophysical Journal*, 815, 117
-

- Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, *The Astrophysical Journal*, 859 [[arXiv]1710.00845]
- Shields, G. A. 1999, *Publications of the Astronomical Society of the Pacific*, 111, 661
- Silverman, J., Ganeshalingam, M., Kong, J., Li, W., & Filippenko, A. 2012, in American Astronomical Society Meeting Abstracts, Vol. 219, American Astronomical Society Meeting Abstracts #219, 126.03
- Smith, M. et al. 2020, *Astrophysical Journal*, 160, 267
- Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757
- Starr, B. M., Luppino, G. A., Cuillandre, J.-C., & Isani, S. 2000, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4008, Optical and IR Telescope Instrumentation and Detectors, ed. M. Iye & A. F. Moorwood, 1022–1033
- Sullivan, M., Conley, A., Howell, D. A., et al. 2010, *Monthly Notices of the Royal Astronomical Society*, 406, 782
- Taubenberger, S. 2017, *The Extremes of Thermonuclear Supernovae*, ed. A. W. Alsabti & P. Murdin, 317
- Taubenberger, S., Hachinger, S., Pignata, G., et al. 2008, *Monthly Notices of the Royal Astronomical Society*, 385, 75
- Taylor, G., Lidman, C., Tucker, B. E., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 504, 4111
- Tripp, R. & Branch, D. 1999, *Astrophysical Journal*, 525, 209
- Tsujikawa, S. 2013, *Classical and Quantum Gravity*, 30, 214003
- Vincenzi, M., Sullivan, M., Möller, A., et al. 2022, *MNRAS* [[arXiv]2111.10382]
- Walker, E. S., Hook, I. M., Sullivan, M., et al. 2011, *Monthly Notices of the Royal Astronomical Society*, 410, 1262
- Wang, X., Li, W., Filippenko, A. V., et al. 2009, *Astrophysical Journal*, 697, 380
- Weinberg, S. 1972, *Gravitation and Cosmology : Principles and Applications of the General Theory of Relativity*
- Westmoquette, M. S., Exter, K. M., Christensen, L., et al. 2009, arXiv e-prints, arXiv:0905.3054
- Wheeler, J. C. & Harkness, R. P. 1990, *Reports on Progress in Physics*, 53, 1467
- Wheeler, J. C. & Levreault, R. 1985, *Astrophysical Journal Letters*, 294, L17
- Yoo, J. & Watanabe, Y. 2012, *International Journal of Modern Physics D*, 21, 1230002
- Zwicky, F. 1937, *Astrophysical Journal*, 86, 217

Résumé

À la fin des années 90, l'utilisation des supernovæ de type Ia (SNe Ia) comme indicateurs de distance a permis la découverte de l'accélération de l'expansion de l'Univers. L'origine de ce fait observationnel est encore mal comprise. Elle est attribuée soit à l'action d'un fluide répulsif de nature inconnue appelé Énergie Noire, soit au fait que la relativité générale est incorrecte aux échelles cosmologiques. La communauté a lancé un programme observationnel ambitieux pour cerner la nature de l'énergie noire qui culminera dans les années 2020, avec quatre relevés dédiés (deux au sol et deux menés depuis l'espace). Le diagramme de Hubble des SNe Ia est un élément majeur de ce programme, la relation distance de luminosité vs. redshift étant particulièrement sensible à l'équation d'état de l'énergie noire $w = p/\rho$. Actuellement, w est mesuré avec une précision de quelques pourcents. La prochaine génération de relevés produira plusieurs dizaines de milliers de SNe, ce qui permettra d'atteindre une précision de l'ordre du pourcent et de détecter une éventuelle variabilité de w avec le *redshift*. Ce résultat est atteignable, d'abord grâce à l'augmentation de la taille des échantillons de SNe Ia et de la qualité de leur suivi mais aussi grâce au contrôle fin des systématiques d'analyses, via le développement d'outils de plus en plus efficaces. Un outil indispensable pour mesurer des distances aux SNe Ia est un modèle empirique spectrophotométrique de l'évolution de leur flux. Il a pour but d'inférer, à partir des courbes de lumière mesurées dans notre référentiel, le flux de chaque supernova *dans son référentiel propre*. Un second but est la mesure d'observables de standardisation corrélées à la luminosité absolue des SNe. L'exploitation de ces corrélations permet de réduire la dispersion des résidus du diagramme de Hubble. Le principal modèle actuellement utilisé par la communauté est le modèle SALT2 (*Spectral Adaptive Light curve Template 2*). Le modèle est construit empiriquement, il est entraîné sur des données de bonne qualité, constituant le lot d'entraînement. Les lots d'entraînements actuels contiennent de l'ordre de mille SNe. Cependant, avec l'avènement des futurs relevés de SNe, le nombre d'objet va augmenter d'un à deux ordres de grandeur.

Durant cette thèse nous avons développé une nouvelle génération de modèle spectrophotométrique, appelée NaCl pour *Nouveaux algorithmes de Courbes de lumière*. Nous sommes repartis de la paramétrisation de SALT2. Nous avons développé un code d'entraînement rapide, écrit dans un langage de programmation moderne (python3), permettant un entraînement sur des dizaines de milliers de SNe en un temps raisonnable. La procédure de minimisation permet une propagation directe des incertitudes affectant les mesures et le modèle, notamment les incertitudes de calibration photométrique de chaque relevé. Les biais et l'efficacité de la méthode ont été étudiés et caractérisés. Nous avons pris soin de séparer la procédure de minimisation et la modélisation du flux des SNe, afin de pouvoir facilement revisiter la paramétrisation de SALT2, et rendre possible l'implémentation de modèles alternatifs.

Abstract

In the late 90's, the use of Type Ia supernovae (SNe Ia) as luminosity distance indicators led to the discovery of the accelerated expansion of the Universe. The origin of the acceleration are not understood yet. It may be due to the presence of a repulsive fluid of unknown nature (Dark Energy) or to the fact that General Relativity is incorrect at cosmological scales. An ambitious observational effort has been started, to understand the nature of Dark Energy. It will culminate in the 2020's with no less than four dedicated surveys (two ground surveys and two space surveys). SNe Ia are a major probe in this program, the redshift-luminosity distance relation being extremely sensitive to the Dark Energy equation of state ($w = p/\rho$). As of today, w is measured with a precision of a few percents. The next generation of surveys will produce $O(10^4)$ SNe, which will allow to reach a precision of $\sim 1\%$ on the measurement of w and $\sim 10\%$ on the potential variations of w with redshift. This result was possible, firstly thanks to the quality and the large number of observations that have been taken since the discovery, but also thanks to the development of more and more efficient analysis tools, which allow a fine control of systematics. An essential tool to measure distances using SNe Ia is an empirical spectrophotometric model of their flux evolution with their phase (time with respect to peak luminosity). The goal of this model is to infer, from the SN light curves measured in observer frame, the peak SN flux in a standard band, *in the SN restframe*. Another goal is to measure, still from the SN lightcurves, standardization observables, which are correlated with SN peak absolute luminosity. Building on these correlations, it is possible to decrease significantly the SN-to-SN dispersion in the Hubble diagram. The model is built empirically. It is trained on good quality data, constituting the training set. They currently contain about a thousand SNe but with the advent of future SNe surveys, the number of objects will increase by one to two orders of magnitude.

In this thesis work, we have developed a new generation spectrophotometric model called NaCl for *Nouveaux algorithmes de Courbes de lumière*, based on the SALT2 parameterization. We have developed a fast model training code, written in a modern programming language (python), that can be trained on tens of thousands of SNe. The minimization procedure allows one to conduct a direct propagation of the model and measurement uncertainties, notably the calibration uncertainties affecting each survey. The biases and the efficiency of the method have been studied and characterized on realistic simulations we have developed. When designing the code, we have been careful to separate the model from the minimization framework, in order to make it easy to explore alternatives to SALT2, in particular, higher dimensionality parametrizations of the SN Ia diversity.