

Création d'un échantillon complet

« La scène disparaît et devient l'un des acteurs. »

Robin GAILLARD

Nous avons vu Chapitre ?? que l'amélioration des mesures de paramètres cosmologiques par le diagramme de HUBBLE nécessite une meilleure précision dans la connaissance astrophysique des SNe Ia, afin, notamment, de permettre la réduction des incertitudes systématiques. À cet effet, une évolution des propriétés intrinsèques des SNe Ia inconnue fausserait ces résultats.

Nous avons vu Chapitre ?? que l'environnement des supernovae avait un impact non négligeable sur leurs caractéristiques mesurées, et notamment leur appartenance à la sous-population « jeune » ou « vieille ». Dans la perspective de mesurer une évolution de la luminosité intrinsèque des SNe Ia, notre recherche se base sur le modèle d'évolution de l'âge moyen de RIGAULT et al. (2020) des SNe Ia et étudie les variations de leur étirement en fonction de l'âge (voir Chapitre ??).

Nous présentons dans ce chapitre l'échantillon sur lequel nous effectuons ces mesures. Nous discutons dans un premier temps des qualités qu'un tel échantillon doit présenter Section I.1, avant de le réaliser et de le présenter Section I.2. La Section I.3 en présente l'augmentation *via* l'implémentation des données de ZTF aux sondages.

Sommaire

I.1 Notion de complétude	2
I.1.1 Stratégies d'observation	2
I.1.2 Biais de MALMQUIST et solution	3
I.2 Échantillon d'étude	3
I.2.1 Confection	3
I.2.2 Présentation	10
I.2.3 Confirmation d'hypothèse	10
I.3 Application à ZTF	14
I.3.1 Redshift limite	14
I.3.2 Présentation	15

I.1 Notion de complétude

Cette thèse repose sur l'étude statistique des propriétés des SNe Ia, et donc en premier lieu sur l'échantillon de données sur lesquelles développer notre raisonnement. Pour qu'il soit intéressant il doit être suffisamment grand, mais également représentatif de la population des SNe Ia. En effet, nous avons vu Chapitre ?? que les SNe Ia sont corrélées avec leur environnement. Comme nous souhaitons étudier toute la zoologie des SNe Ia et non pas une catégorie particulière (par exemple, celles se trouvant uniquement dans des galaxies de masse $M_* > 10 M_\odot$), nous nous intéressons à des données ne comportant pas de biais de la sorte. De même, nous avons vu que pour être de qualité cosmologique, leurs caractéristiques d'étirement et de couleur sont respectivement comprises entre ± 3 et $\pm 0,3$. Pour que cet échantillon soit représentatif de la population des SNe Ia, nous voulons que les données le constituant se rapprochent le plus possible d'un tirage aléatoire de toutes les SNe Ia dans la nature, et non pas, par exemple, d'une seule sous-partie uniquement composée de SNe d'étirement entre ± 2 . Nous définissons alors cet échantillon comme « complet ». Ce concept est largement dépendant de la manière dont les données sont relevées.

I.1.1 Stratégies d'observation

Les supernovae sont des phénomènes transitoires, c'est-à-dire des objets dont le flux lumineux varie dans le temps, et elles sont également brèves et rares : elles durent typiquement quelques semaines et surviennent environ une fois par siècle et par galaxie, toutes SNe confondues, et une fois par millénaire et par galaxie pour les SNe Ia. Leur observation requiert donc des stratégies particulières. Pour déterminer leurs courbes de lumière (??), il est nécessaire d'avoir un champ de mesure suffisamment profond pour ne pas se contenter que de leur luminosité au maximum. Différentes approches peuvent entrer en jeu : les recherches ciblées et les recherches non-ciblées.

Les recherches ciblées consistent à se focaliser sur des amas de galaxies connus en vue d'augmenter la probabilité d'observer des supernovae. Il paraît en effet évident que plus la concentration en étoiles est forte, plus nous nous attendons à une haute probabilité que certaines d'entre elles entament leur fin de vie et leur explosion en supernovae. Cependant, une telle pratique implique une sélection des environnements des SNe et donc un biais sur la nature des données recueillies. Dans le cas des amas de galaxies, l'environnement favorisé sera celui contenant des progéniteurs vieux, dans des galaxies massives avec peu de formation stellaire. Afin d'étudier la potentielle évolution de la population des SNe, il faut réduire au maximum ces biais et favoriser la récolte d'un échantillon représentatif de toute la zoologie des SNe Ia.

Les recherches non-ciblées utilisent de grands champs de caméra pour sonder de larges portions du ciel. Originellement (SCP, PERLMUTTER et al. 1999), leur procédé était d'effectuer une détection photométrique avant d'opérer une identification spectroscopique, confirmant leur caractère de SN Ia ou non, pour finalement décider de programmer ou non un suivi photométrique permettant l'établissement de leur courbe de lumière. Une telle pratique limite les biais mais donne des courbes de lumières pauvres en points de mesure avant le maximum de luminosité, impactant l'ajustement des courbes. Ces méthodes ont

évolué pour devenir des recherches *glissantes* (ASTIER et al. 2006). Elles consistent à balayer régulièrement le ciel en observant un même champ dans un même filtre de manière répétée tous les quelques jours, afin d'à la fois détecter et extraire les courbes de lumières des SNe Ia, même si leur identification est effectuée après leur maximum de luminosité.

I.1.2 Biais de MALMQUIST et solution

De tels sondages ne sont cependant pas exempts d'effets de sélection. En effet, même une recherche glissante s'effectue avec un appareil de mesure ayant une capacité limitée à détecter une source lumineuse : les objets de magnitude apparente plus élevée (luminosité plus faible) que ce seuil de détection ne seront pas inclus. De tels sondages sont dits à magnitude limitée. Or, comme chaque astre voit sa luminosité décroître avec le carré de la distance qui le sépare de l'observation (??), cette limite implique que les astres de magnitude absolue plus élevée seront relevés à de plus grandes distances que les autres, laissant croire qu'à partir d'une certaine distance les objets sont intrinsèquement plus lumineux.

Dans le cadre des SNe Ia dont nous supposons la magnitude absolue similaire, nous pourrions en première approche négliger cet effet. Cependant, comme exposé en Section ??, il a été déterminé que la magnitude absolue des supernovae de type Ia est corrélée avec leur étirement et leur couleur, de telle sorte que les plus faibles soient celles de petit étirement et de couleur rouge. Ainsi, proche du seuil de détection, les SNe Ia ne sont pas sélectionnées de manière homogène, et l'échantillon recueilli sera une sous-population laissant penser qu'avec la distance, les SNe Ia ont en moyenne un plus haut étirement et sont de couleur bleue.

Le cadre de notre étude nécessite un échantillon que nous appelons « volume-limité », pour lequel nous supposons que la population résulte bien d'un tirage aléatoire de ce qui existe dans la nature. Les sondages modernes reposant sur des recherches glissantes, il nous a fallu les réduire pour les utiliser.

I.2 Échantillon d'étude

Nous détaillons dans cette Section la procédure de construction de notre échantillon volume-limité comme expliqué partie I.1.2. Notre étude se base sur les données de la combinaison de sondages Pantheon (SCOLNIC et al. 2018), en remplaçant la combinaison ciblée LOWZ par les données SNfactory dont la sélection est maîtrisée et permettant une étude de sous-population grâce au LsSFR. Les données de HST étant complètes, la confection de notre échantillon se concentre sur les sondages SDSS, PS1 et SNLS ; leur nature non-ciblée et limitée en magnitude permet d'en construire une portion limitée en volume comme décrit Section I.1.

I.2.1 Confection

Nous détaillons ici deux des approches mises en place visant à déterminer la portion des sondages que nous pouvons considérer comme étant limitées en volume.

I.2.1.1 Approche statistique

À partir des données publiées dans [SCOLNIC et al. \(2018\)](#)¹, il est possible de tracer l'histogramme des SNe Ia en fonction du redshift (cf. Sections précédentes, par exemple Figure ??, à gauche). En supposant une densité volumique de supernovae uniforme, chaque intervalle de redshift comprend un volume de plus en plus grand et nous nous attendons donc à observer toujours plus de SNe Ia avec la distance. Nous observons cependant une baisse de ce nombre à partir d'un certain redshift. La chute du nombre de SNe Ia provient de cette limitation du sondage à mesurer la luminosité. Notre première approche a été de se baser sur une étude statistique pour essayer de récupérer la valeur estimée à partir de laquelle chaque sondage s'écarte d'un modèle volumétrique. Le protocole est le suivant :

- Les bornes minimales et maximales des données sont augmentées d'une faible valeur aléatoire afin d'assurer une variation du centre des intervalles. Par exemple, pour SNLS, nous prenons une limite entre 0,06 et 0,12 à gauche et entre 1,10 et 1,15 à droite ;
- Nous choisissons aléatoirement entre 5 et 20 intervalles pour tracer l'histogramme ;
- Nous initialisons un modèle volumétrique $a \times (V(z_2) - V(z_1))$ avec a la densité volumique de SNe Ia, paramètre libre du modèle, auquel nous passons comme donnée les bords des intervalles ;
- Les valeurs du modèle sont comparées aux hauteurs des intervalles de l'histogramme, permettant l'ajustement du modèle par une loi de Poisson cumulée (voir par exemple [AHMED 2015](#)). Pour un intervalle donné de nombre moyen de données λ , la probabilité qu'il y en ait exactement k est, d'après la loi de Poisson :

$$p(k) = \mathcal{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (\text{I.1})$$

La fonction de répartition, ou de distribution cumulative, est donnée par :

$$\mathcal{P}(X \leq x) = \sum_{k=0}^x p(k) = e^{-\lambda} \sum_{k=0}^x \frac{\lambda^k}{k!} \quad (\text{I.2})$$

- Nous choisissons aléatoirement un intervalle maximal après lequel l'ajustement s'arrête, avec un minimum de 3 intervalles (6 dans les cas des Figures de [I.1](#)), 10 fois pour chaque histogramme ;
- Nous sauvons les positions et valeurs de probabilité des intervalles ajustées et créons une interpolation linéaire des résultats ;
- Ces 5 étapes sont répétées 1000 fois et nous calculons la médiane et l'écart type des 10 000 interpolations calculées.

Le modèle volumétrique retenu dans notre analyse est défini par :

$$V(z) = \frac{4\pi}{3} \times d_C^3(z) \quad (\text{I.3})$$

1. https://archive.stsci.edu/hlsps/ps1cosmo/scolnic/data_fitres

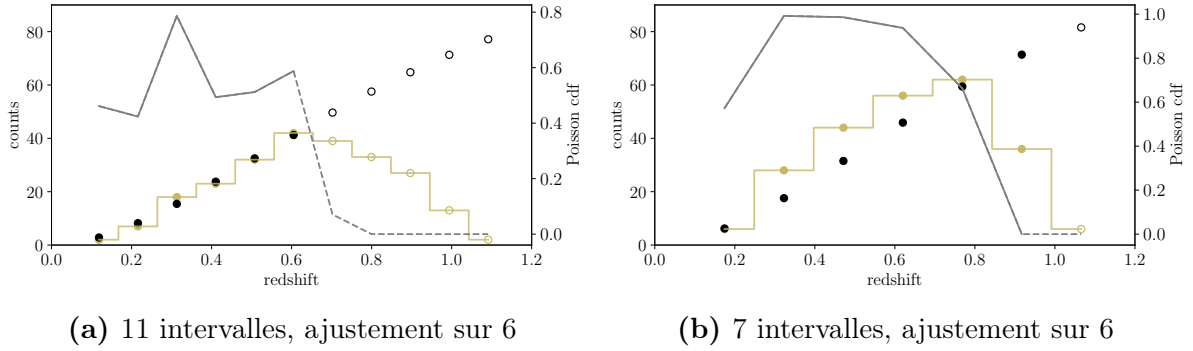


FIGURE I.1 – Exemple d’ajustement statistique pour deux tirages aléatoires d’histogrammes de SNLS.

avec $d_C(z)$ la distance comobile

$$d_C(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')} \quad \text{avec} \quad (I.4)$$

$$E(z) \triangleq \frac{H(z)}{H_0} = [\Omega_R(1+z)^4 + \Omega_M(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda]^{1/2} \quad (I.5)$$

Nous avons choisi la cosmologie issue de la collaboration Planck ([PLANCK COLLABORATION et al. 2020](#)), dont les valeurs sont indiquées Table I.1.

Tableau I.1 – Valeurs des paramètres cosmologiques utilisés pour la détermination statistique du redshift limite des sondages SDSS, SNLS et PS1.

H_0	Ω_R	Ω_M	Ω_k	Ω_Λ
67,74 km Mpc ⁻¹ s ⁻¹	5.389×10^{-5}	0,3075	0	0,6910

Le résultat de ces calculs donne une estimation du redshift à partir duquel chacun des sondages n’a plus la capacité à recueillir toutes les SNe Ia, représentée Figure I.2. En estimant z_{lim} comme étant la valeur à laquelle la médiane des distributions cumulées chute à 0,5 et les erreurs basse et haute à 0,525 et 0,475 respectivement, nous obtenons les valeurs de la Table I.2.

Cette première approche présente une robustesse certaine dans l’établissement des évolutions statistiques en répétant le processus précédent. Cependant, le sens de variation non constant du résultat de SNLS et de PS1 ne permet pas de forte confiance dans la correspondance de ce protocole à l’objectif de cette étude ; de plus, le choix de la valeur de la fonction de répartition à laquelle nous pouvons considérer le sondage complet n’est pas motivée mathématiquement ou physiquement de manière systématique. Cette conclusion nous a amené à une approche combinant à la fois la réalité de la sélection astrophysique instrumentale et les équations de distribution de luminosité de SN Ia avec leurs paramètres x_1 , c et z .

I.2.1.2 Approche analytique

En supposant que ces sondages ont un typage spectroscopique et un suivi photométrique suffisants, ils devraient avoir des effets de sélection de sous-population de SNe Ia négligeables

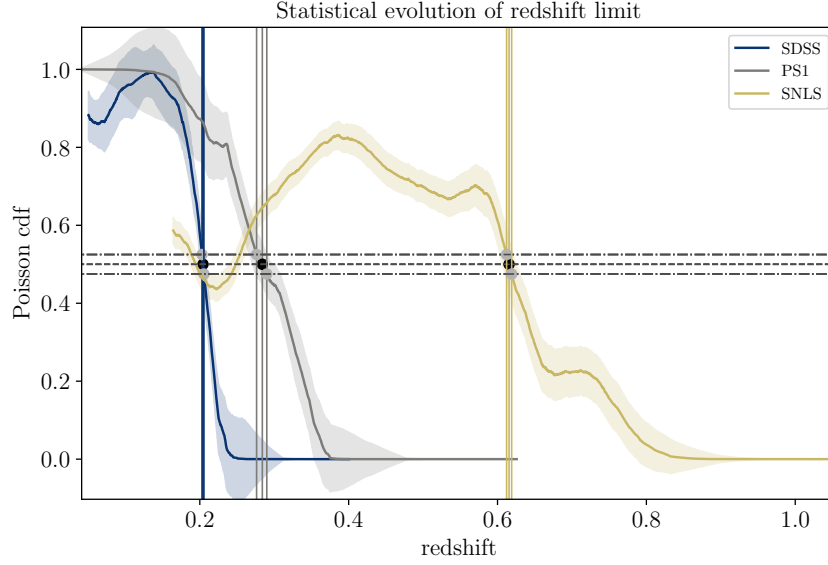


FIGURE I.2 – Résultat graphique de l'évolution médiane de l'étude statistique du redshift limite pour les sondages SDSS, PS1, et SNLS

en deçà d'un certain redshift permettant l'acquisition de toute la zoologie d'étirements et de couleurs. Les données de SNe Ia issues de l'ajustement par **SALT2.4** ne contiennent que des données avec un maximum de $x_1 = \pm 3$ et de $c = \pm 0,3$ (**GUY et al. 2007** ; **BETOULE et al. 2014**, cf Section ??).

La magnitude absolue d'une supernova à son maximum de luminosité est, d'après l'équation ?? :

$$M = M_0 - \alpha x_1 + \beta c$$

avec $M_0 = -19,36$ mag dans le filtre photométrique B de Bessell (**KESSLER et al. 2009** ; **SCOLNIC et al. 2014**), $\alpha = 0,158$ et $\beta = 3,14$ (Table 7, **SCOLNIC et al. 2018**). Nous déterminons cette quantité sur l'ellipse limite des paramètres grâce au paquet **sncosmo**², représentée par un gradient de couleur Figure I.3. Nous trouvons alors que la supernova la moins lumineuse est celle de paramètres $x_1 = -1,65$ et $c = 0,25$ dont le maximum de magnitude absolue standardisée est $M_{\min}^{t_0} = -18,31$ mag.

Cependant, pour établir une courbe de lumière, une supernova doit être observée typiquement au moins 5 jours avant et 1 semaine après son pic de luminosité, donnant une magnitude absolue limite effective d'approximativement $M_{\lim} = -18,00$ mag. En connaissant les magnitudes limites de chaque sondage et avec l'équation reliant le module de distance aux magnitudes observée et absolue

$$\mu(z) = m - M \tag{I.6}$$

nous pouvons déterminer le redshift limite z_{\lim} au-delà duquel la SN Ia la moins lumineuse ne sera pas observée. Nous avons ainsi défini un ensemble de redshifts limites définissant un échantillon fiduciel en choisissant la limite suggérée par cette analyse.

Cependant, cette solution pourrait ne pas être optimale étant donné qu'elle ignore les efficacités de suivi spectroscopiques pour les redshifts en-dessous de z_{\lim} ; c'est pourquoi

2. <https://sncosmo.readthedocs.io/en/stable/>

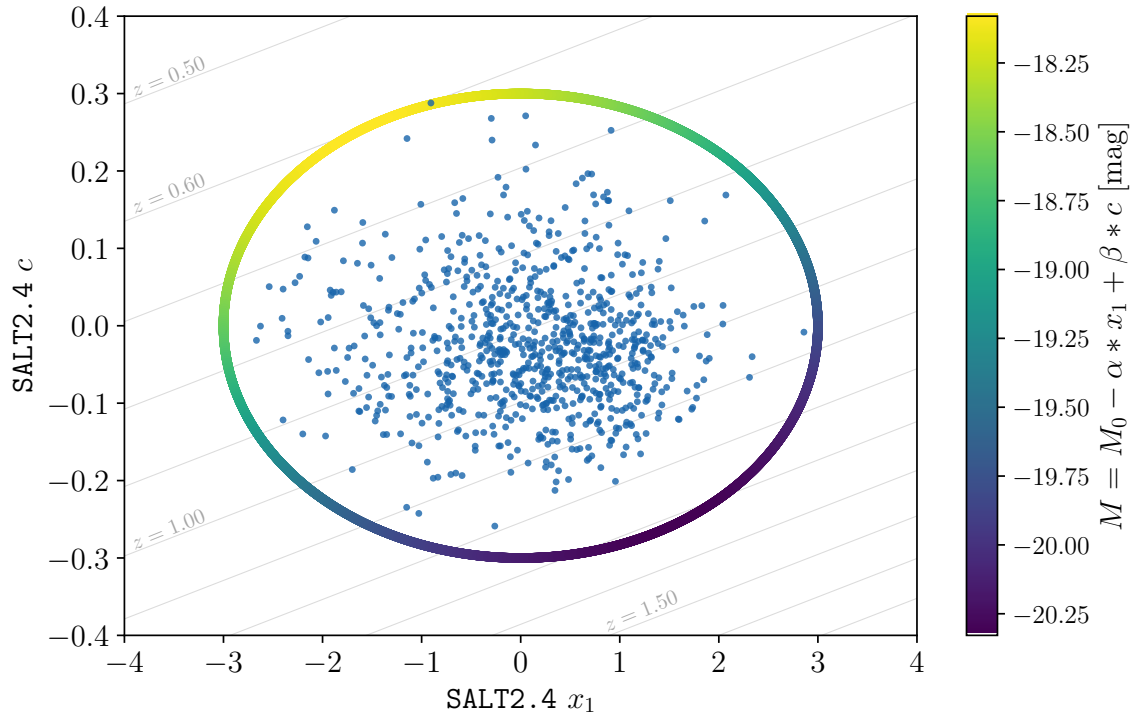


FIGURE I.3 – Distribution des paramètres de courbe de lumière d’étirement (x_1) et de couleur (c) issus d’un ajustement par **SALT2.4** pour les données de SNe Ia des sondages SDSS, PS1 et SNLS combinés du catalogue Pantheon. Chaque supernova est représentée par un point bleu. L’ellipse limite des paramètres ($x_1 = \pm 3, c = \pm 0,3$) est représentée avec un gradient de couleur correspondant à la magnitude absolue standardisée en utilisant les valeurs de **SCOLNIC et al. (2018)** pour les coefficients α et β . Les lignes diagonales grises représentent l’évolution de $m = m_{\text{lim}}$ en fonction de z dans le plan (x_1, c) entre $z = 0,50$ et $z = 1,70$ pour la magnitude limite $m_{\text{lim}} = 24,8$ mag du sondage SNLS.

nous avons également déterminé un autre ensemble de coupes définissant un échantillon « conservatif ». Cet échantillon est plus petit et donc sera statistiquement moins pertinent, mais également moins sujet aux effets de sélection. Ainsi, si l'évolution des propriétés des SNe Ia avec le redshift est encore sondable dans l'échantillon conservatif, il serait encore plus présent dans un échantillon dont l'absence d'effets de sélection est effectuée avec plus de précision que nos coupes en redshift.

Nous présentons Figure I.4 les différentes efficacités spectroscopiques avec le redshift des 4 sondages LOWZ, SDSS, PS1 et SNLS. Nous y observons que les magnitudes limites correspondent également aux limites des capacités spectroscopiques, mais ces seuls critères ne suffisent pas à rendre compte de la qualité limitée en volume d'un sondage. Nous détaillons maintenant les choix pour les sous-échantillons concernés par ces coupes.

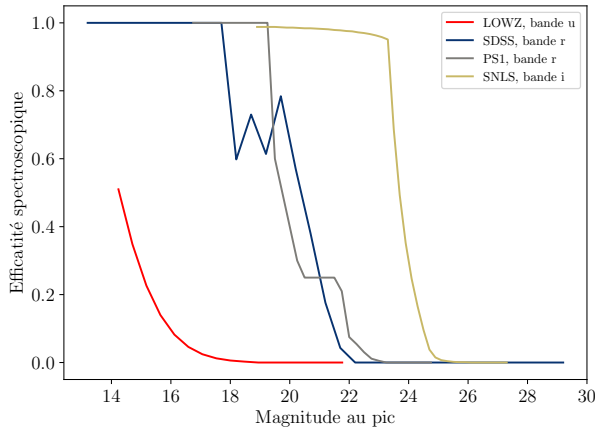


FIGURE I.4 – Comparaison des efficacités spectroscopiques des différents sondages. Figure produite avec les données de la collaboration *Dark Energy Survey* (DES, [ABBOTT et al. 2019](#)) pour LOWZ, de Pantheon ([SCOLNIC et al. 2018](#)) pour PS1, et de Pantheon+ ([POPOVIC et al. 2021](#)) pour SDSS et SNLS.

Pour SNLS dont les supernovae sont typiquement entre $0,4 < z < 0,8$, la bande B de Bessell dans un référentiel au repos correspond approximativement à son filtre i , de magnitude limite à 5σ de 24,8 mag³. Ceci implique $z_{\text{lim}} = 0,60$, en accord avec [NEILL et al. \(2006\)](#) ; [PERRETT et al. \(2010\)](#), et ([CONLEY et al. 2011](#), Section 2.2). D'autre part, la Figure 14 de [PERRETT et al. \(2010\)](#) présentée Figure I.5 suggère une plus basse limite à $z_{\text{lim}} = 0,55$. Nous avons donc choisi $z = 0,60$ et $z = 0,55$ comme redshifts limites de SNLS pour les échantillons fiduciel et conservatif respectivement.

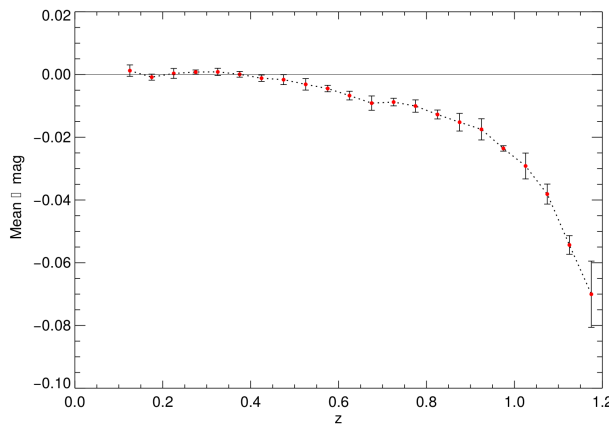


FIGURE I.5 – Biais de MALMQUIST et de sélection spectroscopique moyen en fonction du redshift pour le sondage SNLS d'après des simulations.

Figure de [PERRETT et al. \(2010\)](#).

3. <https://www.cfht.hawaii.edu/Science/CFHTLS/cfhtlsfinalreleaseexecsummary.html>

De la même manière pour PS1 leurs SNe Ia sont entre $0,2 < z < 0,4$; la profondeur à 5σ dans la bande g est de 23,1 mag d'après [REST et al. \(2014\)](#) et mène à $z_{\text{lim}} = 0,31$, en correspondance avec la Figure 6 de [SCOLNIC et al. \(2018\)](#) par exemple, recopiée Figure I.6. De manière conservatrice, cette figure suggère une limite plus prononcée à $z_{\text{lim}} = 0,27$; ces deux valeurs constituent donc les redshifts limites de PS1 pour la partie fiducielle et conservatrice, respectivement, de notre échantillon.

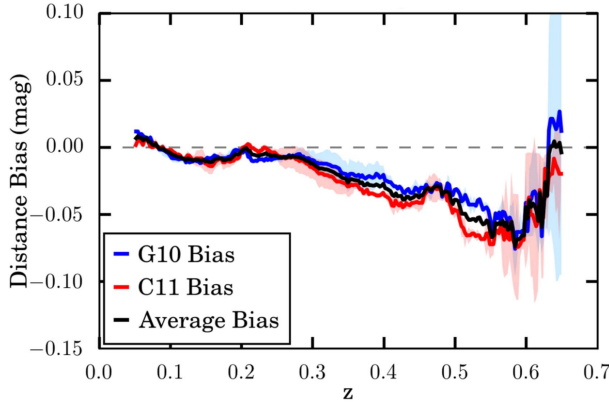


FIGURE I.6 – Biais de MALMQUIST moyen dû aux effets de sélection en fonction du redshift pour le sondage PS1, d'après des simulations et pour deux modèles de dispersion intrinsèques différents et leur moyenne.

Figure de [SCOLNIC et al. \(2018\)](#).

Dans le même intervalle pour SDSS la magnitude limite est de 22,5 mag d'après [DILDAY et al. \(2008\)](#) et [SAKO et al. \(2008\)](#) ; cette valeur impliquerait $z_{\text{lim}} = 0,24$, mais les sondages SDSS se sont confrontés à une limitation dans leurs capacités spectroscopiques. Comme indiqué dans [KESSLER et al. \(2009\)](#) Section 2, les données de la première année de SDSS ont favorisé les SNe Ia de magnitude $r < 20,5$ mag pour identification spectroscopique, ce qui correspondrait à une coupe de redshift à 0,15. Le reste du programme a bénéficié de meilleures ressources spectroscopiques et [KESSLER et al. \(2009\)](#) et [DILDAY et al. \(2008\)](#) font preuve d'une complétude raisonnable jusqu'à $z = 0,2$. La Figure 3 de [CONLEY et al. \(2011\)](#), montrée Figure I.7 et donnant l'évolution du biais de MALMQUIST en fonction du redshift confirme ces hypothèses. En nous basant sur ces faits, nous avons choisi $z_{\text{lim}} = 0,20$ et $z_{\text{lim}} = 0,15$ pour nos échantillons fiduciel et conservatif respectivement.

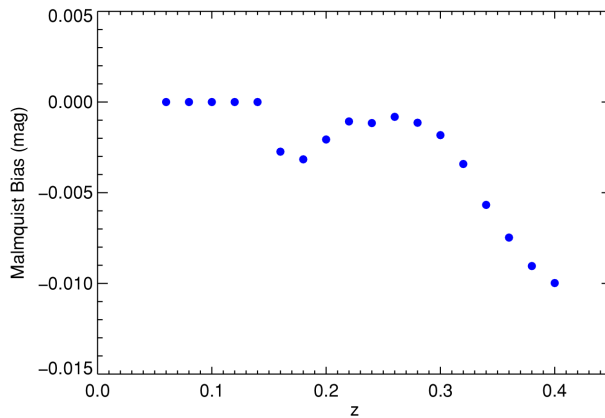


FIGURE I.7 – Biais de MALMQUIST moyen en fonction du redshift pour le sondage SDSS. La forte baisse à $z = 0,15$ est un artefact dû à la discontinuité du modèle d'efficacité spectroscopique et n'a que peu d'effet sur les contraintes cosmologiques.

Figure de [CONLEY et al. \(2011\)](#).

Cette approche est totalement systématique et reproductible, et donne des z_{lim} similaires à l'approche statistique ; cette observation conforte donc les résultats et choix de magnitudes limites, et ce sont ces résultats analytiques que nous avons conservés dans notre étude. La

comparaison des limites par les deux méthodes et le nombre de données conservées avec les limites analytiques dans les cas fiduciels et conservatifs sont indiqués Table I.2.

Tableau I.2 – Composition en SNe Ia de notre échantillon.

Sondage	z_{lim}		N_{SN}
	Statistique	Analytique	
SNf		0,08	114
SDSS	$0,204^{+0,001}_{-0,001}$	0,20 (0,15)	167 (82)
PS1	$0,284^{+0,006}_{-0,008}$	0,31 (0,27)	160 (122)
SNLS	$0,615^{+0,003}_{-0,003}$	0,60 (0,55)	102 (78)
HST	–	–	26
Total	–	–	569 (422)

Notes. L'échantillon et notamment le nombre de SNe utilisées suivent les limites analytiques. Les nombres entre parenthèses correspondent aux limites conservatives.

I.2.2 Présentation

Par rapport aux analyses cosmologiques générales, notre étude impose une forte sélection sur des données déjà soigneusement choisies : seulement 43% (SNLS) à 57% (PS1) de SNe Ia sont conservées. Les distributions en redshift des 3 sondages coupés sont présentées Figure I.8. Nous y observons que les limites sont globalement situées avant le pic de ces histogrammes, suivant la logique guidant cette chute (cf. Section I.1) et confortant également les analyses qui y ont mené. Cette hypothèse est testée Section I.2.3.

En combinant les 5 sondages de notre analyse, nous pouvons tracer leur distribution d'étirement en fonction du redshift. Nous en présentons un graphique ainsi que l'histogramme complet Figure I.9. En supposant l'échantillon affranchi d'effets de sélection, nous pouvons lire sur ce graphique une première idée de l'évolution en redshift que nous supposons issue du changement des propriétés moyennes des SNe Ia avec l'âge de leur environnement. En effet, nous observons que la fraction de SNe Ia présentant un faible étirement, typiquement $x_1 < -1$, semble décroître avec le redshift alors que la population d'étirement > 1 semble toujours peuplée ; à noter qu'ici le redshift est en échelle logarithmique, expliquant le tassement horizontal. Quantitativement, les SNe Ia à haut redshift présentent un plus grand étirement moyen ($0,34 \pm 0,10$ à $z \approx 0,65$) que celles à bas redshift ($-0,17 \pm 0,10$ à $z \approx 0,05$). Cette idée est confirmée dans le chapitre suivant, Section ??.

I.2.3 Confirmation d'hypothèse

Dans la Section précédente, nous avons construit des échantillons limités en volume à partir d'un ensemble d'échantillons limités en magnitude en utilisant des coupures simples de redshift. Cette approche simplifiée est statistiquement sous-optimale, mais devrait suffire pour tester notre question clé, à savoir si l'évolution de l'étirement avec le redshift est compatible avec le modèle de RIGAUD et al. (2020). Cependant, il reste possible qu'une fonction de sélection observationnelle complexe liée aux efficacités de suivi

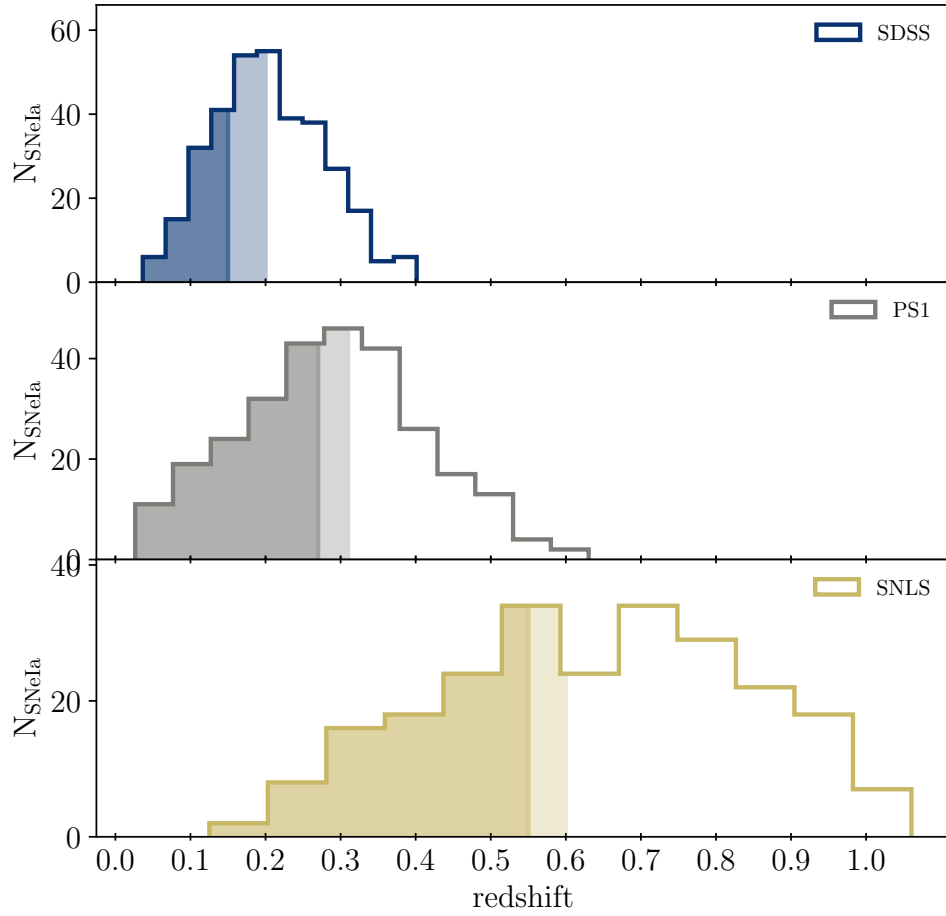


FIGURE I.8 – *De haut en bas* : Histogrammes en redshift des SNe Ia des sondages SDSS, PS1 et SNLS (données de Pantheon, [SCOLNIC et al. 2018](#)). Les parties colorées représentent les distributions de SNe Ia conservées dans notre analyse, considérées exemptes d’effets de sélection observationnels (cf. Section I.2.1.2). Les couleurs foncées (claires) représentent les limites conservatives (fiducielles) de nos coupes de sélection indiquées dans la Table I.2.

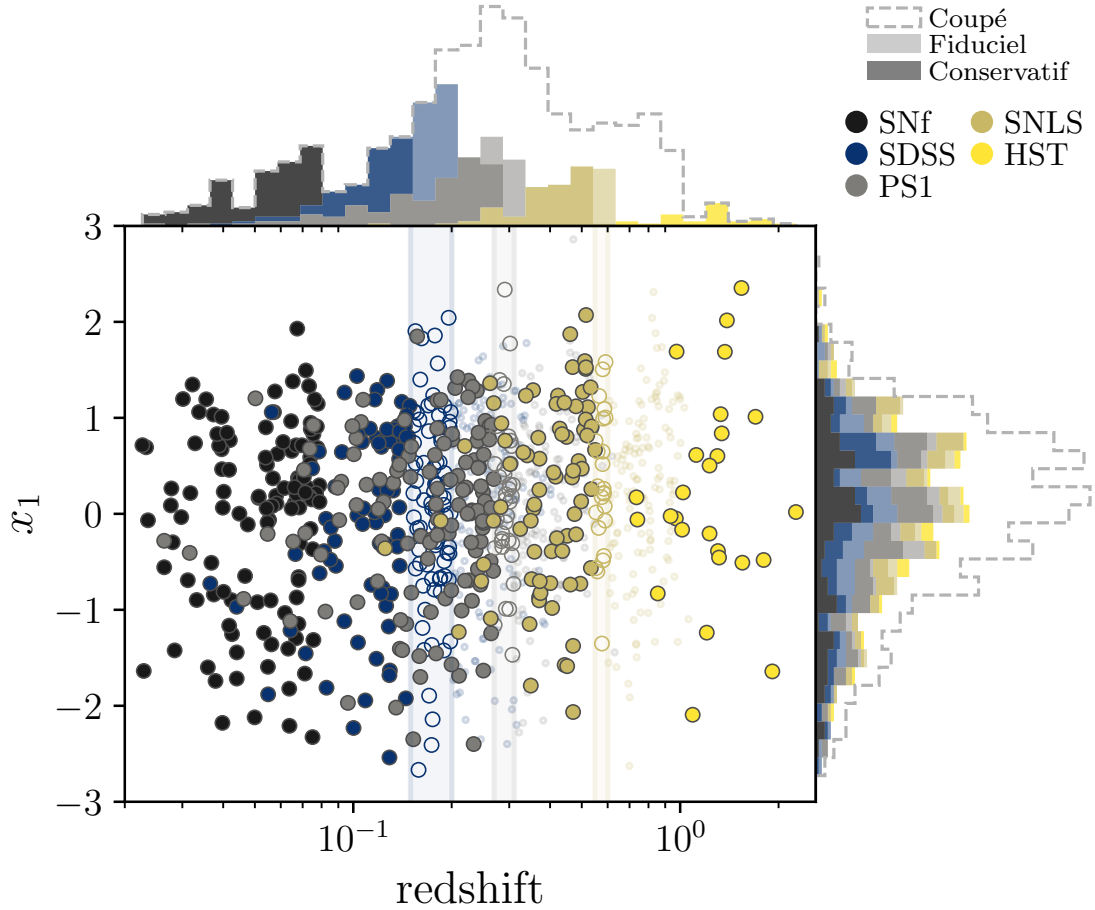


FIGURE I.9 – *En bas* : étirement des courbes de lumière ajustées avec SALT2.4 en fonction du redshift en échelle logarithmique pour chaque sondage de cette analyse (cf. légende). Les points pleins, creux et transparents correspondent aux échantillons conservatif, fiduciel et total, respectivement. *En haut* : histogrammes en redshift superposés, en sombre, clair et pointillé pour les échantillons conservatif, fiduciel et total, respectivement (cf. légende). À *droite* : histogrammes en étirement superposés. Notre hypothèse de travail se base sur le dépeuplement des étirements < -1 avec le redshift, que nous pouvons apercevoir dans les histogrammes de droite : relativement aux données $x_1 > -1$, cette partie se peuple moins.

spectroscopique en deçà de nos coupures fiducielles (voire conservatives) puisse encore affecter notre échantillon, le rendant non entièrement limité en volume ; cela fausserait alors notre conclusion sur la dérive astrophysique de la population SNe Ia.

Pour tester l'existence d'éventuels biais de sélection observationnels dans notre échantillon, nous avons comparé les distributions d'étirement et de couleur des SNe Ia provenant de différents ensembles de données dont les plages de redshifts se chevauchent : ces distributions devraient être similaires si elles reflètent la même population mère sous-jacente. Nous notons que la plage de redshift doit être suffisamment étroite pour que toute dérive soit négligeable.

Les deux échantillons qui se chevauchent le plus en termes de redshift sont PS1 et SDSS dans la plage de redshift $0,10 < z < 0,20$ (voir Figure I.9). Ce sous-échantillon est constitué des 146 SNe Ia à l'extrémité haute des redshifts de SDSS et est donc le plus susceptible d'être affecté par des effets de sélection observationnels résiduels (voir la discussion correspondante dans la Section I.1.2). Sur cette même plage de redshift, PS1 compte 52 SNe Ia qui se trouvent dans les tranches de redshift les plus basses et qui sont donc peu susceptibles d'être affectées par un problème de sélection observationnelle. Afin d'identifier les incohérences potentielles entre les sous-échantillons PS1 et SDSS, la Figure I.10 (panneaux supérieurs) compare la distribution des étirements et des couleurs de ces deux études. Les valeurs p du test de similarité de KOLMOGOROV-SMIRNOV (KS) qui en résultent ($p > 10\%$) n'indiquent aucune raison de croire que les deux distributions ne sont pas issues d'une même population sous-jacente, ce qui aurait été le cas si l'une ou l'autre des distributions était affectée par des effets de sélections, en accord avec l'impression visuelle de la Figure I.10.

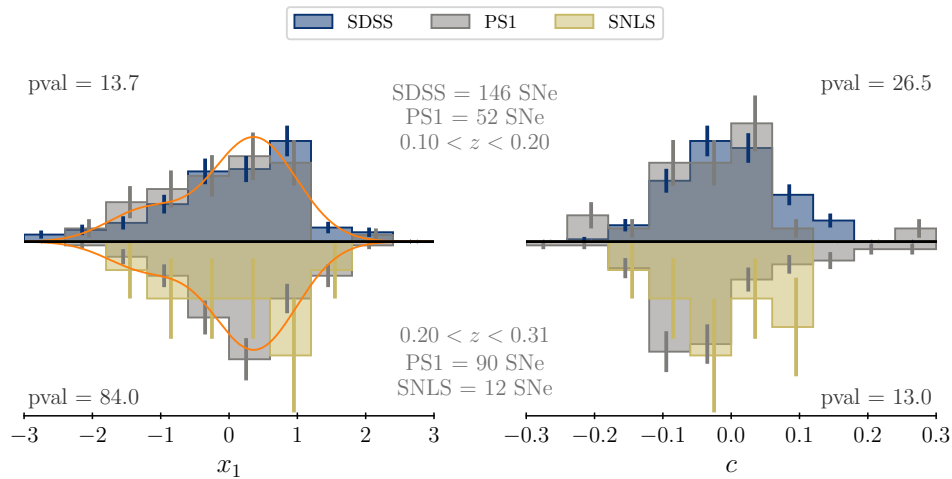


FIGURE I.10 – Histogrammes de distribution d'étirement (gauche) et de couleur (droite) de différents relevés se chevauchant en redshift. *Vers le haut* : SDSS et PS1 dans la plage de redshift $0,10 < z < 0,20$. *Vers le bas* : PS1 et SNLS dans la gamme de redshift $0,20 < z < 0,31$. Les barres d'erreur représentent le bruit de Poisson. Notre modèle d'évolution en redshift (défini plus loin, Section ??) est illustré en orange à la valeur moyenne des plages de redshifts, 0,15 et 0,25 respectivement. Les valeurs p du test de KOLMOGOROV-SMIRNOV sont indiquées en haut (en bas) de chaque panneau et ne montrent aucun signe que les distributions d'étirement et de couleur de SDSS et PS1 (PS1 et SNLS) ne sont pas tirées des mêmes distributions sous-jacentes.

Nous avons effectué une analyse similaire pour PS1 et SNLS sur la plage de redshift

$0,20 < z < 0,31$ (Figure I.10, panneaux inférieurs), où la même conclusion peut être tirée : il n'y a pas de signe significatif de divergence dans les distributions d'étirement et de couleur entre les extrémités basse et haute de nos échantillons fiduciels de SNLS et PS1, respectivement. Néanmoins, la petite taille de l'ensemble de données SNLS à $z < 0,31$ (12 SNe Ia contre 90 pour PS1) limite la sensibilité de ce test, et seule une forte déviation serait perceptible. L'extension de la plage de redshift à $0,20 < z < 0,40$ (bien que nous n'ayons pas de données PS1 au-dessus de 0,3) permet d'augmenter le sous-échantillon SNLS à 31, mais la valeur p de l'étirement reste élevée (34%).

Nous soulignons enfin que la couleur des SNe Ia est plus sujette aux effets de sélection observationnelle que l'étirement, comme l'illustre la Figure I.3 ; voir également la Figure 3 de KESSLER et SCOLNIC (2017), par exemple. Par conséquent, comme la comparaison des distributions de couleurs ne montre aucune indication significative d'un effet de sélection observationnelle résiduel, cela renforce notre affirmation selon laquelle nos critères de sélection simples basés sur le redshift sont suffisants pour construire les échantillons complets de SNe Ia nécessaires pour tester l'évolution du redshift de la distribution d'étirement.

I.3 Application à ZTF

Au moment de rédiger cette thèse, la deuxième publication de données de ZTF est en cours. Ainsi, bien qu'il ne fasse pas partie de cette étude de manière stable, nous nous intéressons à son intégration dans notre travail. C'est pourquoi nous présentons ici notre confection de la partie limitée de ZTF.

I.3.1 Redshift limite

Sa magnitude limite est donnée à 20,4 mag⁴, ce qui se convertirait à un $z_{\text{lim}} = 0,11$, mais la SEDm est indiquée avoir une classification correcte entre $[18,5 - 19]$ mag (FREMLING et al. 2020), ce qui implique des limites $z_{\text{lim}} = [0,044 - 0,055]$ respectivement. Ces résultats sont confortés par une étude du sondage effectuée par AMENOUCHE et al. (*en préparation*), utilisant le module Python `simsurvey`⁵ de ZTF (FEINDT et al. 2019). Initialement utilisé pour permettre le pronostic des résultats du sondage, l'utilisation de journaux d'observation réels indique que l'échantillon devrait être limité en volume jusqu'à $z \approx 0,06$. Ainsi, ce sont les valeurs $z_{\text{lim}} = 0,044$ et $z_{\text{lim}} = 0,055$ qui constituent nos limites conservative et fiducielle pour ZTF, respectivement. Le Tableau I.3 résume les informations de coupe pour ZTF.

Tableau I.3 – Composition en SNe Ia pour notre échantillon ZTF.

Sondage	z_{lim}	N_{SN}
ZTF	0,055 (0,044)	638 (393)
Total	–	1379 (987)

Notes. Le nombre total inclut les données de la section précédente.

4. <https://www.ztf.caltech.edu/ztf-camera.html>

5. <https://github.com/ZwickyTransientFacility/simsurvey>

I.3.2 Présentation

Dans le cas de ZTF, la coupe en redshift est encore plus sévère que pour les autres sondages ; c'est en effet 28% des données qui sont conservées. Il reste cependant que sa partie fiducielle est à elle seule plus peuplée que la totalité de notre échantillon de base. L'inclusion de ces données augmentera donc la statistique et la robustesse de notre étude d'un facteur non-négligeable. La Figure I.11 en montre la distribution des données d'étirement en fonction du redshift.

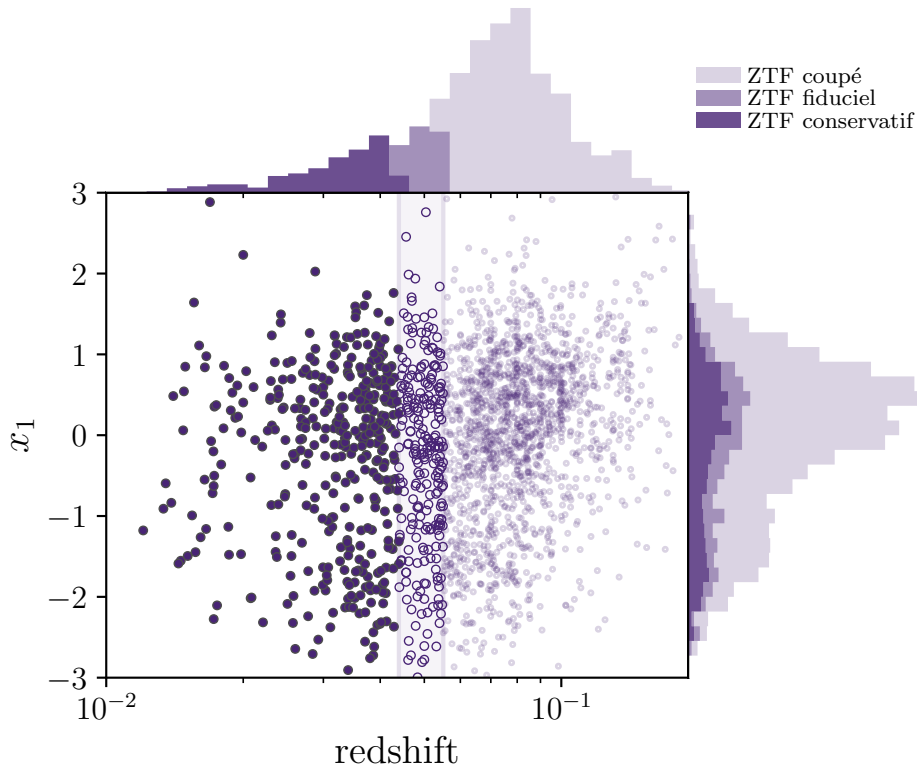


FIGURE I.11 – *En bas* : étirement des courbes de lumière ajustées avec **SALT2.4** en fonction du redshift en échelle logarithmique pour ZTF (cf. légende). Les points pleins, creux et transparents correspondent aux parties conservative, fiducielle et totale respectivement. *En haut* : histogrammes en redshift superposés, en sombre, clair et très clair pour les parties conservative, fiducielle et totale respectivement (cf. légende). *À droite* : histogrammes en étirements superposés, même légende.

Sur cette figure, nous pouvons d'ores et déjà observer la présence d'un sursaut à basse valeur dans la distribution des étirements ($x_1 \approx -1.5$). Sa présence est en accord avec l'étude de la distribution d'étirement du chapitre suivant. Ces données étant arrivées après l'étude en question, sa présence en conforte les résultats ; nous en discutons Section ??.

Au sein de ce sondage, la dérive de l'âge avec le redshift est considérée comme négligeable d'après l'Équation ?? dont l'évolution est montrée Figure ??. Pour voir une éventuelle variation de la population à petit étirement avec le redshift, nous présentons Figure I.12 l'échantillon de base combiné aux données de ZTF. La présence de nombreuses données à $z \lesssim 0,05$ dû à ZTF écrase les données des autres échantillons s'étalant jusqu'à $z = 2.26$ pour

la représentation logarithmique, mais permet une meilleure visualisation de la potentielle évolution de l'étirement avec le redshift grâce aux histogrammes de la partie droite.

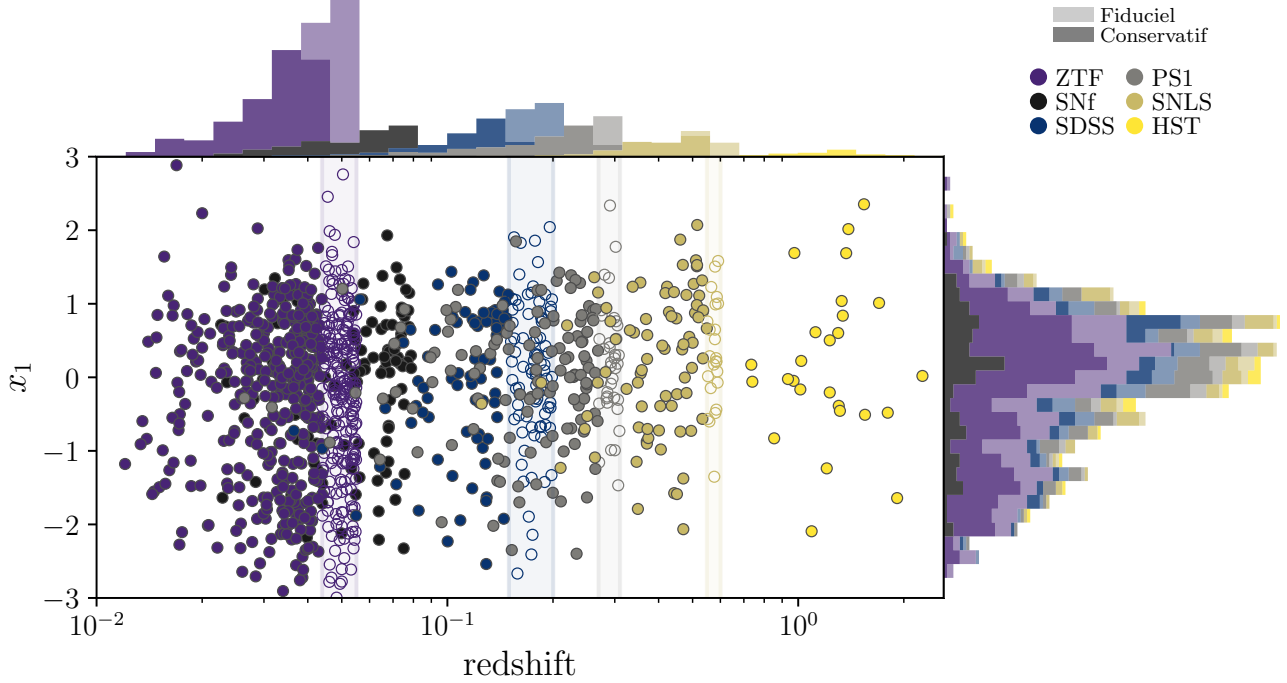


FIGURE I.12 – *En bas* : étirement des courbes de lumière ajustées avec SALT2.4 en fonction du redshift en échelle logarithmique pour l'intégrité des sondages apparaissant dans l'étude de la dérive de l'étirement avec le redshift (cf. légende). Les points pleins et creux correspondent aux parties conservatrice et fiducielle respectivement. *En haut* : histogrammes en redshift superposés, en sombre et clair pour les parties conservatrice et fiducielle respectivement (cf. légende). *À droite* : histogrammes en étirements superposés, même légende.

Figures

I.1 Exemple d'ajustement statistique pour deux tirages aléatoires d'histogrammes de SNLS.	5
I.2 Évolution médiane du redshift limite des sondages SDSS, PS1 et SNLS par approche statistique	6
I.3 Distribution et limite des paramètres de courbe de lumière d'étirement (x_1) et de couleur (c) des sondages SDSS, PS1 et SNLS combinés du catalogue Pantheon	7
I.4 Comparaison des efficacités spectroscopiques des différents sondages	8
I.5 Biais de MALMQUIST moyen en fonction du redshift pour le sondage SNLS	8
I.6 Biais de MALMQUIST moyen en fonction du redshift pour le sondage PS1	9
I.7 Biais de MALMQUIST moyen en fonction du redshift pour le sondage SDSS	9
I.8 Histogrammes des sondages coupés pour notre étude	11
I.9 Présentation des données d'étirement en fonction du redshift pour l'échantillon complet	12
I.10 Histogrammes de test de similarité de KOLMOGOROV-SMIRNOV entre les sondages SDSS et PS1 d'une part, PS1 et SNLS d'autre part, en étirement et en couleur	13
I.11 Présentation des données d'étirement en fonction du redshift pour ZTF	15
I.12 Présentation des données d'étirement en fonction du redshift pour l'échantillon de base combiné aux données de ZTF	16

Tableaux

I.1 Valeurs des paramètres cosmologiques utilisés pour la détermination statistique du redshift limite des sondages SDSS, SNLS et PS1	5
I.2 Composition en SNe Ia de notre échantillon.	10
I.3 Composition en SNe Ia pour notre échantillon ZTF.	14

Bibliographie

- ABBOTT T. M. C., ALLAM S., ANDERSEN P. et al. 2019, « First Cosmology Results using Type Ia Supernovae from the Dark Energy Survey: Constraints on Cosmological Parameters », [ApJ](#), **872**, L30 ↑ [Page 8](#)
- AHMED S. N. 2015, in *Physics and Engineering of Radiation Detection (Second Edition)*, second edition edn., ed. S. N. Ahmed (Elsevier), 541–593 ↑ [Page 4](#)
- ASTIER P., GUY J., REGNAULT N. et al. 2006, « The Supernova Legacy Survey: measurement of Ω_M , Ω_Λ and w from the first year data set », [A&A](#), **447**, 31 ↑ [Page 3](#)
- BETOULE M., KESSLER R., GUY J. et al. 2014, « Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples », [A&A](#), **568**, A22 ↑ [Page 6](#)
- CONLEY A., GUY J., SULLIVAN M. et al. 2011, « Supernova Constraints and Systematic Uncertainties from the First Three Years of the Supernova Legacy Survey », [ApJS](#), **192**, 1 ↑ [Page 8](#), ↑ [Page 9](#)
- DILDAY B., KESSLER R., FRIEMAN J. A. et al. 2008, « A Measurement of the Rate of Type Ia Supernovae at Redshift $z \approx 0.1$ from the First Season of the SDSS-II Supernova Survey », [ApJ](#), **682**, 262 ↑ [Page 9](#)
- FEINDT U., NORDIN J., RIGAULT M. et al. 2019, « simsurvey: estimating transient discovery rates for the Zwicky transient facility », [J. Cosmology Astropart. Phys.](#), **2019**, 005 ↑ [Page 14](#)
- FREMLING C., MILLER A. A., SHARMA Y. et al. 2020, « The Zwicky Transient Facility Bright Transient Survey. I. Spectroscopic Classification and the Redshift Completeness of Local Galaxy Catalogs », [ApJ](#), **895**, 32 ↑ [Page 14](#)
- GUY J., ASTIER P., BAUMONT S. et al. 2007, « SALT2: using distant supernovae to improve the use of type Ia supernovae as distance indicators », [A&A](#), **466**, 11 ↑ [Page 6](#)
- KESSLER R., BERNSTEIN J. P., CINABRO D. et al. 2009, « SNANA: A Public Software Package for Supernova Analysis », [PASP](#), **121**, 1028 ↑ [Page 6](#), ↑ [Page 9](#)
- KESSLER R. et SCOLNIC D. 2017, « Correcting Type Ia Supernova Distances for Selection Biases and Contamination in Photometrically Identified Samples », [ApJ](#), **836**, 56 ↑ [Page 14](#)
- NEILL J. D., SULLIVAN M., BALAM D. et al. 2006, « The Type Ia Supernova Rate at $z \sim 0.5$ from the Supernova Legacy Survey », [AJ](#), **132**, 1126 ↑ [Page 8](#)
- PERLMUTTER S., ALDERING G., GOLDBABER G. et al. 1999, « Measurements of Ω and Λ from 42 High-Redshift Supernovae », [ApJ](#), **517**, 565 ↑ [Page 2](#)

- PERRETT K., BALAM D., SULLIVAN M. et al. 2010, « Real-time Analysis and Selection Biases in the Supernova Legacy Survey », [AJ](#), **140**, 518 ↑ [Page 8](#)
- PLANCK COLLABORATION, AGHANIM N., AKRAMI Y. et al. 2020, « Planck 2018 results. VI. Cosmological parameters », [A&A](#), **641**, A6 ↑ [Page 5](#)
- POPOVIC B., BROUT D., KESSLER R. et SCOLNIC D. 2021, « The Pantheon+ Analysis: Forward-Modeling the Dust and Intrinsic Colour Distributions of Type Ia Supernovae, and Quantifying their Impact on Cosmological Inferences », [arXiv e-prints](#), [arXiv:2112.04456](#) ↑ [Page 8](#)
- REST A., SCOLNIC D., FOLEY R. J. et al. 2014, « Cosmological Constraints from Measurements of Type Ia Supernovae Discovered during the First 1.5 yr of the Pan-STARRS1 Survey », [ApJ](#), **795**, 44 ↑ [Page 9](#)
- RIGAULT M., BRINNEL V., ALDERING G. et al. 2020, « Strong dependence of Type Ia supernova standardization on the local specific star formation rate », [A&A](#), **644**, A176 ↑ [Page 1](#), ↑ [Page 10](#)
- SAKO M., BASSETT B., BECKER A. et al. 2008, « The Sloan Digital Sky Survey-II Supernova Survey: Search Algorithm and Follow-up Observations », [AJ](#), **135**, 348 ↑ [Page 9](#)
- SCOLNIC D., REST A., RIESS A. et al. 2014, « Systematic Uncertainties Associated with the Cosmological Analysis of the First Pan-STARRS1 Type Ia Supernova Sample », [ApJ](#), **795**, 45 ↑ [Page 6](#)
- SCOLNIC D. M., JONES D. O., REST A. et al. 2018, « The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample », [ApJ](#), **859**, 101 ↑ [Page 3](#), ↑ [Page 4](#), ↑ [Page 6](#), ↑ [Page 7](#), ↑ [Page 8](#), ↑ [Page 9](#), ↑ [Page 11](#)