

# Impact sur la cosmologie : simulations

*« The Answer to the Great Question... Of Life,  
the Universe and Everything... Is... Forty-two. »*

Douglas ADAMS, *H2G2*

Nous avons vu dans le chapitre précédent la manière dont **SNANA** permettait de traiter les biais et corrélations environnementales dans le calcul des paramètres cosmologiques, et ainsi pourquoi son utilisation dans notre thèse était pertinente.

Dans ce chapitre, nous présentons les simulations que nous avons effectuées avec le logiciel. Dans un premier lieu, nous discutons des différentes corrélations que nous avons testées *via* l'utilisation de **HOSTLIB** (Section I.1) et de la confection des nôtres (Section I.2). Par la suite, nous sondons la qualité de chacun de ces modèles à décrire les données (Section I.3) pour finalement regarder l'impact de ces différentes hypothèses sur la valeur calculée de  $w$  (Section I.4). Nous concluons Section I.5.

## Sommaire

<b>I.1 Présentation des HOSTLIB . . . . .</b>	<b>2</b>
I.1.1 Étirement et couleur globales : SK . . . . .	2
I.1.2 Étirement et couleur selon la masse : BP . . . . .	3
I.1.3 Étirement selon l'âge : NN . . . . .	3
I.1.4 Étirement et marche de magnitude selon l'âge : NR . . . . .	3
<b>I.2 Confection des HOSTLIB NN et NR . . . . .</b>	<b>4</b>
I.2.1 Modélisation du lien entre masse et redshift . . . . .	4
I.2.2 Comparaison aux données . . . . .	6
I.2.3 Sélection des modèles . . . . .	6
I.2.4 Génération des HOSTLIB . . . . .	7
I.2.5 Implémentation . . . . .	10
Résumé . . . . .	11
<b>I.3 Comparaison des données simulées aux données réelles . . . . .</b>	<b>12</b>
I.3.1 Accord entre les données : analyse uni-dimensionnelle . . . . .	13
I.3.2 Accord entre les données : analyse bi-dimensionnelle . . . . .	15
<b>I.4 Impact sur la cosmologie . . . . .</b>	<b>18</b>
I.4.1 Résultats de standardisation . . . . .	18
I.4.2 Résultats de cosmologie . . . . .	21
I.4.3 Systématiques dues au choix du modèle de masse . . . . .	23
<b>I.5 Conclusion . . . . .</b>	<b>25</b>

## I.1 Présentation des HOSTLIB

Dans sa forme la plus générale, une HOSTLIB ne possède pas de valeurs liées à des paramètres de SNe Ia (comme l'étirement ou la couleur) ; en effet, elle sert originellement à utiliser le redshift photométrique de la galaxie hôte comme valeur antérieure dans l'ajustement du redshift de la SN et à ajouter du bruit à la SN simulée (voir Chapitre ??). Avant d'intégrer notre modèle à SNANA, il nous a fallu reproduire les approches d'autres groupes utilisant le logiciel. Nous avons choisi pour cela les études de SCOLNIC et KESSLER (2016), ci-après SK, et de POPOVIC et al. (2021), ci-après BP.

### I.1.1 Étirement et couleur globales : SK

Dans leurs travaux, SK n'incluent pas de lien d'étirement ou de couleur avec les propriétés de la galaxie hôte mais uniquement une marche de magnitude en fonction de sa masse  $M_*$ . Celle-ci est inclut dans les WEIGHTMAP des sondages, et est de 0,05 mag. Le tirage des paramètres  $x_1$  et  $c$  se font alors depuis des distributions asymétriques Gaussiennes, une par sondage simulé, décrites par :

$$P(p) = \begin{cases} e^{-\frac{|p - \mu|^2}{\sigma_-^2}} & \text{si } p \leq \mu \\ e^{-\frac{|p - \mu|^2}{\sigma_+^2}} & \text{si } p > \mu \end{cases} \quad (\text{I.1})$$

avec  $p = x_1$  ou  $c$ . Les valeurs des paramètres sont indiqués Tableau I.1.

**Tableau I.1** – Paramètres des distributions sous-jacentes d'étirement et de couleur desquelles sont générées les SNe Ia dans notre reproduction du travail de SK.

Sondage	$x_1$			$c$		
	$\mu$	$\sigma_-$	$\sigma_+$	$\mu$	$\sigma_-$	$\sigma_+$
PS1	0,604	1,029	0,363	-0,077	0,029	0,12
SDSS	1,141	1,653	0,100	-0,038	0,048	0,079
SNLS	0,964	1,232	0,282	-0,065	0,044	0,12
LOWZ	–	–	–	-0,055	0,023	0,015

**Notes.** Les valeurs viennent du Tableau 1 de SK, sauf pour LOWZ dont la distribution d'étirement est une double Gaussienne d'après SCOLNIC et al. (2018).

Nous avons cependant utilisé les valeurs de paramètres de SCOLNIC et al. (2018) pour la distribution d'étirement de LOWZ, qui est alors décrite par une combinaison de deux Gaussiennes dont une asymétrique, telle que :

$$P(x_1) = A_1 \times \begin{cases} e^{-\frac{|x_1 - \mu_1|^2}{\sigma_{-,1}^2}} & \text{si } x_1 \leq \mu_1 \\ e^{-\frac{|x_1 - \mu_1|^2}{\sigma_{+,1}^2}} & \text{si } x_1 > \mu_1 \end{cases} + A_2 \times e^{-\frac{|x_1 - \mu_2|^2}{\sigma_2^2}} \quad (\text{I.2})$$

Les valeurs sont indiquées Tableau I.2 avec le rapport d'amplitude  $a = \frac{A_1}{A_2}$ .

**Tableau I.2** – Paramètres de la distribution sous-jacente d’étirement pour l’échantillon LOWZ dans notre reproduction de l’étude de SK.

Sondage	$x_1$					
	$\mu_1$	$\sigma_{-,1}$	$\sigma_{+,1}$	$a$	$\mu_2$	$\sigma_2$
LOWZ	0,55	1,0	0,45	0,55	-1,5	0,5

**Notes.** Les caractéristiques sont celles reportées dans l’annexe C de SCOLNIC et al. (2018), mais les valeurs y étant erronées nous avons utilisé celles de l’équipe directement.

### I.1.2 Étirement et couleur selon la masse : BP

D’un autre côté, BP définissent des distributions mères Gaussiennes asymétriques pour  $x_1$  et  $c$  selon la masse de la galaxie hôte. Ceci est effectué en découplant les données des sondages en intervalles selon  $M_*$ ; dans chacun de ces intervalles sont déterminés les paramètres des Gaussiennes asymétriques, puis à chaque entrée de la HOSTLIB sont sélectionnés des paramètres d’étirement et de couleur selon la valeur de la masse de la galaxie hôte. Ainsi, par rapport à SK, ces HOSTLIB présentent 2 colonnes supplémentaires, une pour  $x_1$  et une pour  $c$ , attribuant à chaque entrée une valeur de ces paramètres à associer à la SN simulée.

Ce procédé est réalisé pour LOWZ d’une part, menant à une HOSTLIB que nous appelons « BP\_lowz », et pour la combinaison des sondages DES, SDSS, PS1 et SNLS d’autre part, menant à une HOSTLIB que nous nommons « BP\_highz ». Les valeurs des paramètres correspondants sont disponibles dans l’annexe A2 de BP.

Ce sont ces HOSTLIB qui constituent la base de notre étude ; en réalité, les HOSTLIB SK sont celles de BP où nous avons retiré le tirage des colonnes  $x_1$  et  $c$ .

### I.1.3 Étirement selon l’âge : NN

Notre approche des corrélations entre environnement et supernova est une variation forte par rapport aux précédentes implémentations, puisqu’elle résulte d’une modélisation prospective plutôt que purement phénoménologique. Pour notre étude, nous avons besoin de relier l’étirement attribué à la SN avec l’âge de son environnement, en correspondance avec nos travaux précédents (NICOLAS et al. 2021, ci-après NN). Nous augmentons donc les HOSTLIB en faisant correspondre un âge à chaque entrée des tables ; ainsi, par rapport aux HOSTLIB BP, nous avons donc une colonne indiquant si la SN est jeune ou vieille, et nommons ces HOSTLIB « NN ». La réalisation de cette HOSTLIB est présentée dans la Section I.2.

### I.1.4 Étirement et marche de magnitude selon l’âge : NR

Comme nous l’avons vu précédemment, l’âge des SNe Ia a une double implication : celle de l’évolution de la distribution sous-jacente de l’étirement avec le redshift (Chapitre ??), mais aussi une marche de magnitude de  $\gamma_{\text{env}} = 0,13$  mag entre les SNe Ia jeunes et vieilles (Chapitre ??, RIGAULT et al. 2020). Nous avons implémenté cette valeur à la place de la marche de magnitude selon  $M_*$  incluse dans les WEIGHTMAP des sondages via l’ajout d’une colonne donnant une variation de magnitude de  $\pm 0,065$  mag aux HOSTLIB NN : ces

nouvelles HOSTLIB se nomment « NR » pour « NICOLAS RIGAULT », et présentent l'exacte même colonne d'étirement que les NN.

## I.2 Confection des HOSTLIB NN et NR

Afin de simuler des SNe Ia avec notre modèle, que ce soit pour NN ou NR, nous avons besoin que les propriétés des galaxies hôtes suivent les distributions de ce qui a été observé par les différents sondages simulés. Bien que nous soutenions que le LsSFR est un meilleur traceur de l'environnement d'une SN (BRIDAY et al. 2022), la plupart des sondages caractérisent les galaxies avec leur masse stellaire. Étant donné que le LsSFR repose sur la masse de la galaxie hôte<sup>1</sup>, nous ne pouvons pas uniquement nous baser sur la valeur du redshift  $z$  d'une entrée des HOSTLIB BP pour y assigner un étirement. Nous attendons effectivement que les galaxies vers  $M_* \approx 12$  ne contiennent des SNe Ia vieilles alors que les galaxies de  $M \gtrsim 7$  n'en contiennent que des jeunes.

Par conséquent, afin de comparer les implications de notre modélisation basée sur le LsSFR avec ce que les autres sondages ont observé, nous avons dû modéliser les masses des galaxies hôtes par rapport au LsSFR ; nous avons pour cela utilisé le même échantillon du Chapitre ?? que pour la modélisation de l'étirement du Chapitre ???. Cependant, nous soulignons que cette étude n'a pas pour volonté de décrire l'évolution des masses des galaxies hôtes (qui sont des propriétés globales) avec le LsSFR d'une supernova (étant une propriété intrinsèque de celles-ci) : son utilité est d'associer de manière cohérente un âge à une SN caractérisée par un certain redshift et par une masse de galaxie hôte.

### I.2.1 Modélisation du lien entre masse et redshift

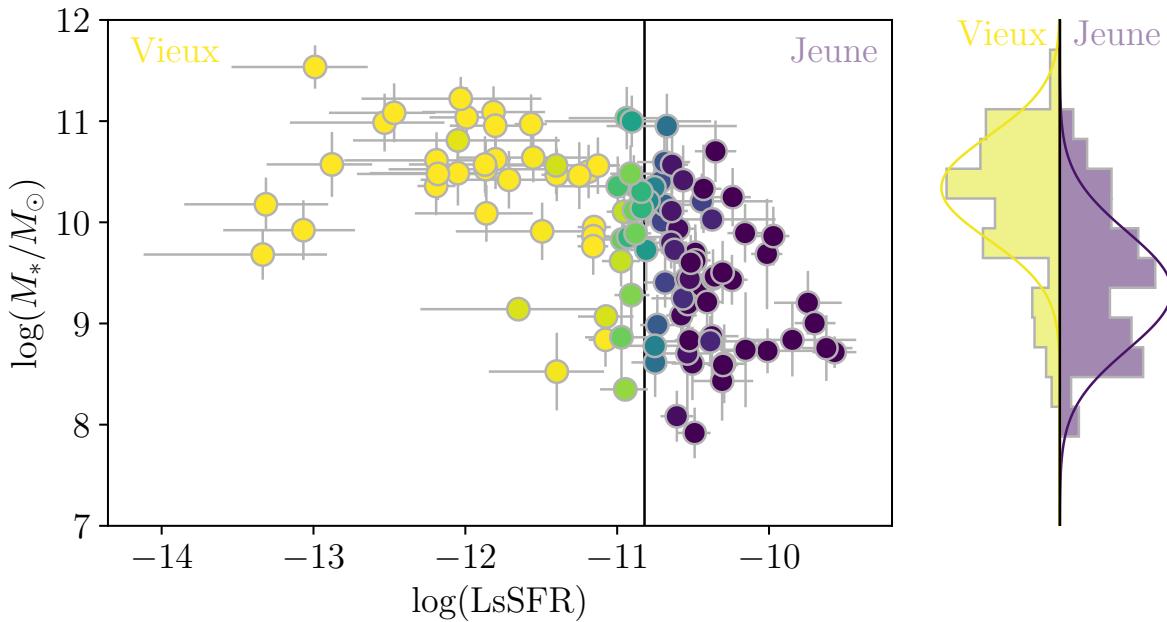
De la même manière que dans le Chapitre ??, nous utilisons le LsSFR comme traceur de l'âge d'une SN, mais cette fois sur les estimations de masse du sondage SNf. Ensuite nous modélisons les populations jeune et vieille par une série de paramétrisations différentes et choisissons celle qui a le plus faible AIC. Cependant, les masses SNf ont été calculées à l'aide de l'Équation 8 de TAYLOR et al. (2011) (voir RIGAULT et al. (2020)) alors que d'autres études du catalogue Pantheon utilisent différentes techniques d'estimation de la masse qui pourraient donner des valeurs de sortie différentes pour une même galaxie.

L'estimation de TAYLOR utilise la magnitude AB absolue en bande  $i$  d'une galaxie,  $M_i$ . Elle est déduite de la magnitude apparente  $m_i$  connaissant le redshift de la galaxie mais suppose que la bande  $i$  observée est proche de celle du référentiel de repos, ce qui est vrai pour les redshift de SNf qui sont inférieurs à  $z \approx 0,05$ . Les relevés de l'échantillon Pantheon sont à des redshifts plus élevés et ont utilisé un ajustement des distributions spectrales d'énergie (en anglais SED pour *spectral energy distributions*) des galaxies pour éviter les corrections K dans cette procédure.

Nous avons appliqué la même analyse à l'échantillon SNf ; les données ainsi calculées sont nommées « SEDSNf » par la suite, et l'échantillon fiduciel utilisant SEDSNf est nommé « SED fiduciel ». À cause de l'absence de 4 galaxies hôtes dans les catalogues de données nécessaires à ce calcul, l'échantillon SEDSNf est réduit à 110 données, au lieu de 114 pour SNf. Nous indiquons également que nous nous limitons pour cette étude

---

1. pour rappel :  $\text{sSFR} = \frac{\text{SFR}}{M_*}$



**FIGURE I.1 – Principal :** masses des galaxies hôtes ( $M_*$ ) ajustées par SED en fonction du LsSFR pour les SNe de SNfactory. La couleur correspond à la probabilité  $p_y$  que la SN Ia soit jeune, c'est-à-dire qu'elle ait  $\log \text{LsSFR} \geq -10,82$  (voir RIGAULT et al. 2020, et Chapitre ??). À droite : histogramme pondéré par  $p_y$  des étirements des SNe, ainsi que le modèle sélectionné ajusté ; les contributions des populations jeune et âgée sont indiquées en violet et jaune, respectivement.

aux galaxies de  $M_* > 10^7 M_\odot$ , étant donné que les valeurs inférieures servent à indiquer qu'aucune galaxie hôte n'a été définie ; l'échantillon fiduciel tombe alors à 548 données, et SED Fiduciel à 544.

Nous avons donc réalisé cette étude de l'évolution des distributions sous-jacentes de masse avec le LsSFR en utilisant :

**SNf** Uniquement les données de SNf (masses issues d'un calcul par l'Équation de TAYLOR) ;

**SEDSNf** Uniquement les données de SNf avec les masses issues d'un ajustement par SED ;

**Fiduciel** Toutes les données de notre échantillon fiduciel ;

**SED fiduciel** Toutes les données de notre échantillon fiduciel avec les masses SEDSNf.

D'après la forme des histogrammes de la Figure I.1, nous avons implémentés différentes modélisations. Cette étude étant annexe à la simulation par SNANA, nous ne présentons que les plus pertinentes et omettons les modélisations n'ayant pas d'intérêt physique ou mathématique, c'est-à-dire les modélisations constantes avec le redshift (notamment les Gaussienne simple et Gaussienne asymétrique pure) et les modèles ne convergeant pas. Ainsi, nous présentons les modèles suivants :

- « Howell » d'après HOWELL et al. (2007), avec une Gaussienne simple pour chacune des populations jeune et âgée (voir Chapitre ??) ;
- « Howell+asym » où la population jeune est une simple Gaussienne et la population vieille est une Gaussienne asymétrique ;

- « Howell asym » où les deux populations jeune et âgée sont asymétriques.

### I.2.2 Comparaison aux données

Chacun de ces modèles a été ajusté aux différents échantillons, et nous en présentons maintenant les résultats. La procédure d'ajustement est celle de la Section 3 de [NN](#), selon la présence de LsSFR dans chaque sous-échantillon. Nous définissons de même que précédemment

$$-2 \ln(L) = -2 \sum_i \ln \mathcal{P} \left( x_1^i \mid \vec{\theta}; dx_1^i, y^i \right). \quad (\text{I.3})$$

et nous utilisons le critère d'information d'AKAIKE (AIC, [BURNHAM 2004](#)) pour comparer la capacité de chaque modèle à décrire correctement les données en pénalisant l'ajout de paramètres libres tel que :

$$\text{AIC} = -2 \ln(L) + 2k, \quad (\text{I.4})$$

ce qui permet d'éviter le sur-ajustement. Les résultats sont présentés Tableau [I.3](#).

**Tableau I.3** – Comparaison de la capacité relative de chaque modèle à décrire les données selon l'échantillon d'ajustement.

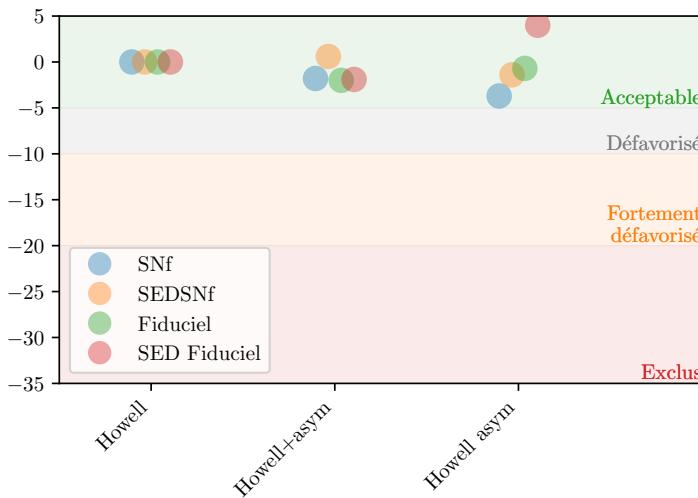
Échantillon	N <sub>SNe Ia</sub>	Howell ( $k = 4$ )			Howell+asym ( $k = 5$ )			Howell asym ( $k = 6$ )		
		-2 ln( $L$ )	AIC	ΔAIC	-2 ln( $L$ )	AIC	ΔAIC	-2 ln( $L$ )	AIC	ΔAIC
SNf	114	230,0	238,0	–	229,8	239,8	-1,8	229,7	241,7	-3,7
SEDSNF	110	223,9	231,9	–	221,4	231,4	0,6	221,3	233,3	-1,4
Fiduciel	544	1534,3	1542,3	–	1534,3	1544,3	-2,0	1531,0	1543,0	-0,7
SED Fiduciel	548	1546,6	1554,6	–	1546,5	1556,5	-1,9	1538,7	1550,7	4,0

**Notes.** Pour chaque modèle considéré, nous indiquons son nombre de paramètres libres  $k$ , et pour chaque échantillon étudié son  $-2 \ln(L)$  (voir Équation ??), son AIC et la différence d'AIC ( $\Delta\text{AIC}$ ) entre ce modèle et le modèle Howell, choisi comme référence car présentant l'AIC le plus faible pour 6 comparaisons sur 8.

Après calcul, le modèle Howell est celui qui se détache le plus, étant celui de plus petit AIC pour 6 modèles sur 8, et est celui représenté sur la Figure [I.1](#); cependant tous les modèles sont considérés comme étant de bonnes représentations des données. Nous présentons Figure [I.2](#) une illustration des résultats du tableau précédent, et Figure [I.3](#) les représentations graphiques des modèles implémentés variant en redshift.

### I.2.3 Sélection des modèles

Avec la multitude de modèles possibles pour établir nos [HOSTLIB](#), nous avons dû effectuer une sélection. Étant donné que notre but est d'associer de manière cohérente un âge de SN définie par un redshift et une masse de galaxie hôte, une caractéristique primordiale au modèle choisi est d'avoir une évolution de la fraction de jeunes SNe Ia physiquement cohérente avec les observations; nous nous attendons notamment à ce que la fraction de jeunes étoiles soit  $\approx 1$  pour les  $M_* \gtrsim 10^7 M_\odot$ , diminue progressivement jusqu'à  $\approx 50\%$  pour  $M_* \approx 10^{10} M_\odot$  et continue sa progression vers 0 pour  $M_* > 10^{10} M_\odot$ ; en effet,



**FIGURE I.2** –  $\Delta\text{AIC}$  entre le modèle Howell et les autres modèles (voir Tableau I.3). Tous les modèles sont dérivants. Les marqueurs bleus, orange, verts, rouges montrent les résultats lorsque l’analyse est effectuée sur l’échantillon SNf, SEDSNf, fiduciel, fiduciel avec SEDSNf, respectivement (voir légende). Les bandes de couleur illustrent la validité des modèles, d’acceptable ( $\Delta\text{AIC} > -5$ ) à exclu ( $\Delta\text{AIC} < -20$ ). En suivant ces valeurs d’AIC, tous les modèles sont compatibles entre eux.

la position de la marche de magnitude basée sur la masse est à  $M_* = 10^{10}\text{M}_\odot$  et cette limite constitue un bon indicateur de l’âge d’une SN Ia d’après BRIDAY et al. (2022).

Nous avons étudié cette évolution pour les différents modèles implémentés, dont les résultats sont présentés Figure I.4.

Nous trouvons alors que tous les modèles finissent par présenter une remontée de la fraction de jeunes étoiles quand la masse  $M_* > 10^{11}\text{M}_\odot$ , sauf le modèle Howell ajusté sur l’échantillon SEDSNf. Cela provient de l’incertitude des courbes Gaussiennes des sous-populations jeunes étant bien plus larges que celles des sous-populations vieilles, donnant pour les masses élevées un rapport de probabilité en faveur des jeunes SNe Ia. Ceci ne correspondant pas à une réalité physique, nous rejetons tous les modèles Howell+asym et Howell asym de notre étude à partir de ces résultats. Parmi les modèles Howell, seul celui ajusté sur SNf passe en effet par 50% à  $M_* = 10^{10}\text{M}_\odot$ .

Nous conservons ainsi les modèles suivants :

- 1) Le modèle Howell ajusté sur SEDSNf;
- 2) Le modèle Howell ajusté sur SNf;

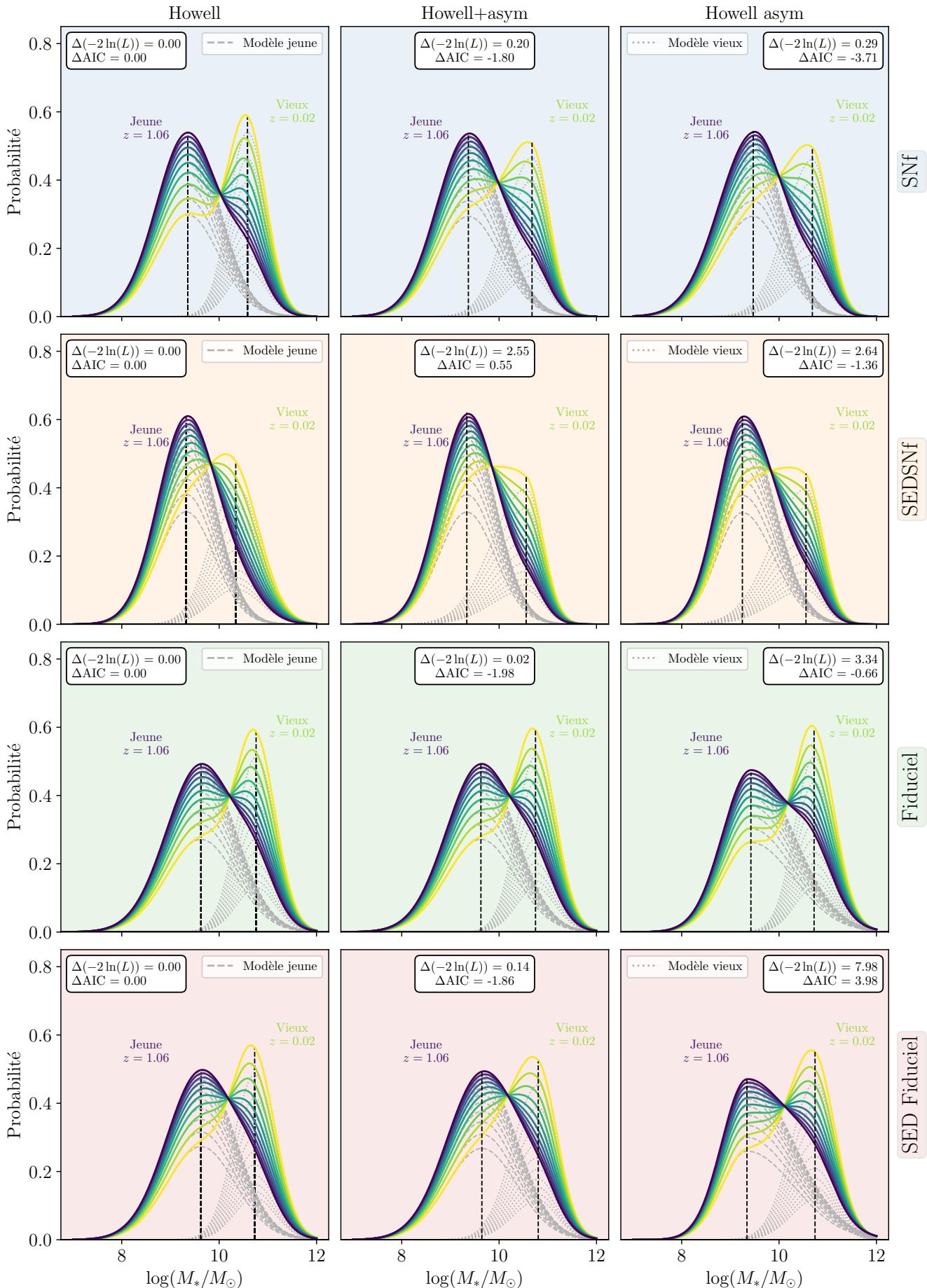
et pour reproduire artificiellement la descente de la fraction de jeunes étoiles en fonction de la masse attendue, nous avons également :

- 3) « SNfsupp » (« suppressed », « réprimé ») : le modèle Howell ajusté sur SNf, mais pour lequel les objets de  $M_* > 10^{11}\text{M}_\odot$  sont automatiquement associés à des SNe Ia âgées.

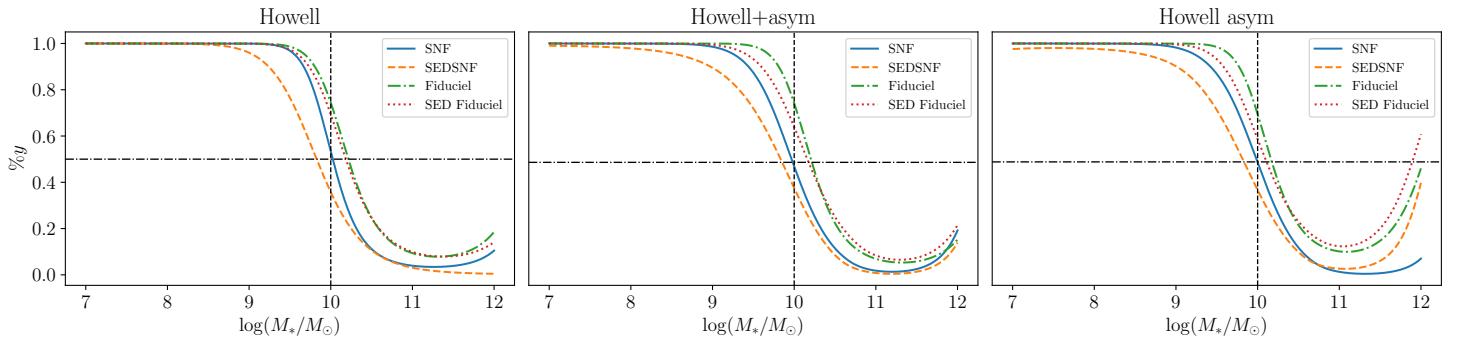
Nous prenons le modèle SNfsupp comme référence ; les implications du choix de modélisation de masse est discuté Section I.4.3. Les valeurs des paramètres des modèles SEDSNf et SNf sont indiquées Tableau I.4.

#### I.2.4 Génération des HOSTLIB

Avec les modélisations de la masse et de l’étirement en fonction du redshift, nous pouvons à présent lire les entrées des HOSTLIB BP, et à partir d’un redshift générer une



**FIGURE I.3** — Modèles implémentés et testés dans l'étude de l'évolution de la masse avec le redshift. Les modèles Howell, Howell+asym et Howell asym sont tracés dans la colonne de gauche, du milieu et de droite, respectivement. Les échantillons sur lesquels ils sont ajustés correspondent aux lignes et à la couleur de fond du graphique : SNf (bleu), SEDSNf (orange), fiduciel (vert), fiduciel avec SEDSNf (rouge); ce sont les mêmes couleurs que dans la Figure I.2. Les quantités  $\Delta(-2\ln(L))$  et  $\Delta\text{AIC}$  par rapport au modèle Howell de chaque ligne sont indiquées pour chaque modèle figure. Nous avons tracé dix réalisations des modèles selon la valeur du redshift moyen considéré, de la valeur la plus basse de notre échantillon ( $z = 0.02$ ) à la valeur maximale des données totales (sans coupe en redshift) de SNLS ( $z = 1.06$ ) représentés en couleur allant du jaune (bas redshift, plus vieil environnement) au violet (haut redshift, environnement jeune) et les distributions des populations jeune et vieille constituant le modèle total sont en gris pointillé et fin pointillé, respectivement. Nous y retrouvons l'information que tous les modèles sont compatibles en tant que bonnes représentations des données par rapport au modèle de base.



**FIGURE I.4** – Comparaison de la prédiction de l’évolution de la fraction de jeunes SNe Ia ( $\%y$ ) en fonction de la masse de la galaxie hôte ( $M_*$ ) pour chaque modèle et selon chaque échantillon utilisé pour l’ajustement. Alors que le sens de variation devrait être constant, pratiquement tous les modèles finissent par remonter après  $M_* \approx 10^{11} M_\odot$ , sauf le modèle Howell ajusté sur SEDSNf. Nous excluons les modèles Howell+asym et Howell asym par ce critère.

**Tableau I.4** – Valeurs des paramètres issus des meilleurs ajustement du modèle Howell sur les échantillons SNf et SEDSNf.

Échantillon	$\mu_y$	$\sigma_y$	$\mu_o$	$\sigma_o$
SNf	$9.36 \pm 0.06$	$0.64 \pm 0.04$	$10.58 \pm 0.04$	$0.38 \pm 0.04$
SEDSNf	$9.32 \pm 0.07$	$0.58 \pm 0.05$	$10.34 \pm 0.07$	$0.51 \pm 0.06$

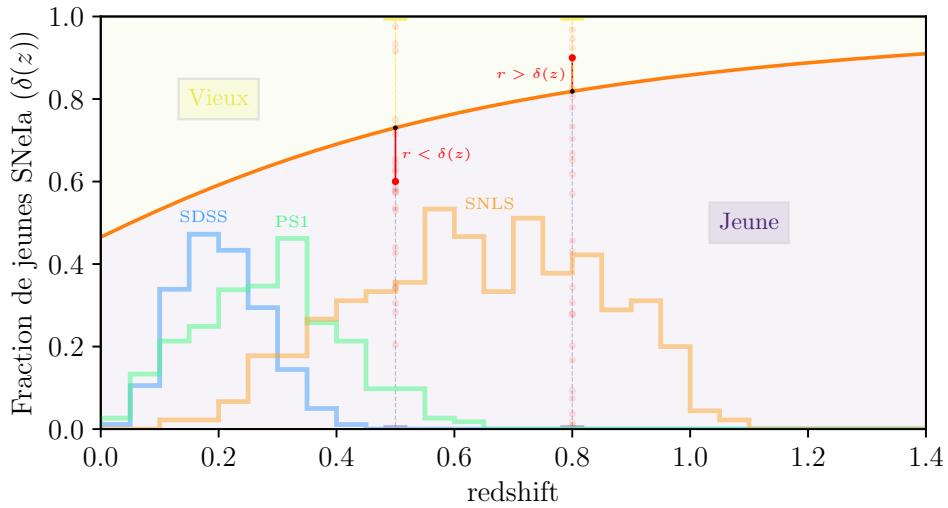
liste de masses et d’étirements. Cela nous permettra ensuite de faire correspondre la masse de la HOSTLIB avec celles de la liste générée, et d’attribuer une valeur d’étirement qui remplacera celle de BP.

Cette étape est réalisée avec le module Python `SNprop`<sup>2</sup>. Ce processus prend la fraction attendue de jeunes étoiles en utilisant  $\delta(z)$  donnée Équation ???. Il assigne une qualité « jeune » ( $LsSFR = 1$ ) ou « vieille » ( $LsSFR = 0$ ) au tirage qui va suivre en prenant un nombre aléatoire  $r$  entre 0 et 1 et en le comparant à la valeur de la fraction susmentionnée. Si  $r < \delta(z)$ , alors la SN simulée sera jeune et inversement. Plus  $z$  augmente et plus  $\delta(z)$  augmente, et donc plus la probabilité d’être assigné jeune augmente. Ceci est présenté Figure I.5.

Cette étape est réalisée 1000 fois pour chaque redshift de la HOSTLIB, donnant une table de redshift, âge (0 ou 1), masse et étirement de 1000 entrées, puis une correspondance est effectuée entre toutes les masses tirées et la masse de la HOSTLIB pour trouver celle qui en est la plus proche. Nous prenons alors la valeur d’étirement associée et remplaçons celle de la HOSTLIB. Au même moment, nous entrons la valeur de l’âge (0 ou 1) dans une nouvelle colonne ; ceci conclut la confection des HOSTLIB NN.

Les HOSTLIB NR possèdent un autre colonne supplémentaire, où à chaque valeur d’âge est associée une valeur de variation de magnitude, de +0,065 mag pour les jeunes (moins lumineuses) et de -0,065 mag pour les vieilles (plus lumineuses), qui remplacent le valeurs de marche de magnitude basées sur la masse implémentées dans les autres approches et qui sont associées au WEIGHTMAP.

2. <https://github.com/MickaelRigault/snprop>

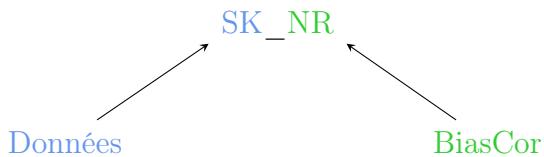


**FIGURE I.5** – Représentation du choix de l’âge d’une SN et de l’assignation de masse et d’étirement en fonction du redshift du module Python `SNprop`<sup>2</sup>. *Orange* : fraction estimée de jeunes SNe Ia en fonction du redshift. *Histogrammes* : nombres de SNe Ia des 3 sondages principaux de l’échantillon Pantheon (SCOLNICK et al. 2018) (pas à l’échelle). *Lignes verticales* : pour chaque  $z$  de la HOSTLIB, un nombre aléatoire  $r$  entre 0 et 1 est tiré : s’il est supérieur (inférieur) à  $\delta(z)$  à ce redshift, alors la SN sera assignée vieille (jeune) et les valeurs de masse et d’étirement générées seront tirées des distributions sous-jacentes vieilles (jeunes) des paramètres correspondants.

## I.2.5 Implémentation

Nous pouvons implémenter différentes manières d’effectuer ces simulations, que nous appelons « types ». Une approche serait de simuler 100 fois des échantillons de la taille de l’échantillon de Pantheon ( $\approx 1000$ ) et de combiner les résultats, permettant ainsi d’avoir des incertitudes statistiques réalistes. Bien que nous ayons entamé la réalisation d’une telle approche, le plus simple et moins coûteux en temps a été de simuler un échantillon d’une taille conséquente ( $\approx 13\,000$ ) avec un BiasCor 50 fois plus grand, donnant une idée de l’incertitude systématique due aux différents modèles de corrélations (SK, BP, NN, NR).

Pour quantifier cela, nous conservons les données simulées et les échantillons BiasCor associés de chaque modélisation, afin d’utiliser les données d’un modèle et de les corriger avec le BiasCor d’un autre : l’idée derrière cette pratique est de connaître le potentiel biais dû au fait de méconnaître la physique réelle qui régit les propriétés intrinsèques des SNe Ia. Pour les distinguer, nous les nommons de la manière suivante :



Ainsi, « SK\_NR » décrit un échantillon dont les données ont été générées en supposant les modèles de corrélations de SK et corrigées avec des données générées en supposant les modèles de corrélations dus à l’âge (NR). Lorsque les données et BiasCor sont les mêmes, nous ne mentionnons pas quel est le BiasCor.

Pour comparer de manière cohérente les données simulées aux données réelles, il faut que le ratio des données de chaque sondage de l'échantillon simulé corresponde au ratio des données de chaque sondage de l'échantillon réel. Ceci s'effectue *via* un paramètre appelé `NGEN`, décrivant le nombre d'années de sondage simulé. Il permet de contrôler plus ou moins précisément le nombre de SNe Ia simulées, puisque chaque sondage a sa propre efficacité spectroscopique qui, à chaque simulation, opère une sélection des données conservées (voir Chapitre ??). Notamment, puisque l'efficacité spectroscopique de l'échantillon LOWZ est particulièrement faible, il nécessite un grand `NGEN` dans nos fichiers de configurations. De plus, la correction par `BBC` réduit l'échantillon en ne conservant que les données qui sont dans un intervalle de `BiasCor` avec suffisamment de points pour avoir une valeur de correction. Nous indiquons dans le Tableau I.5 le nombre de données pour les données réelles et pour nos simulations, exprimées en pourcentages de l'échantillon Pantheon.

**Tableau I.5** – Nombre de données après l'ajustement par `BBC` et après l'échantillonnage nécessaire à la reproduction des ratio observés dans l'échantillon Pantheon ([SCOLNICK et al. 2018](#)).

Données	BiasCor	Après BBC				
		Total (%)	LOWZ	SDSS	PS1	SNLS
Pantheon		1022	172	335	279	236
SK	SK	13333 (13.05)	13.64	7.29	19.63	13.00
	BP	12847 (12.57)	13.31	7.00	19.24	12.06
	NN	12898 (12.62)	13.03	7.01	19.40	12.27
	NR	12898 (12.62)	13.03	7.01	19.40	12.27
BP	SK	12316 (12.05)	10.50	6.71	18.10	13.61
	BP	12462 (12.19)	10.59	6.77	18.66	13.42
	NN	12397 (12.13)	10.02	6.76	18.68	13.55
	NR	12397 (12.13)	10.02	6.76	18.68	13.55
NN	SK	12439 (12.17)	12.59	6.54	17.87	13.12
	BP	12478 (12.21)	12.51	6.57	18.33	12.75
	NN	12787 (12.51)	13.09	6.61	18.73	13.11
	NR	12787 (12.51)	13.09	6.61	18.73	13.11
NR	SK	12461 (12.19)	13.01	6.60	17.89	12.81
	BP	12475 (12.21)	12.88	6.62	18.32	12.41
	NN	12798 (12.52)	13.49	6.70	18.72	12.76
	NR	12798 (12.52)	13.49	6.70	18.72	12.76

**Notes.** Les pourcentages sont indiqués par rapport à la taille de l'échantillon Pantheon, voir première ligne. Les sous-échantillons simulés sont indiqués en pourcentages directement.

## Résumé

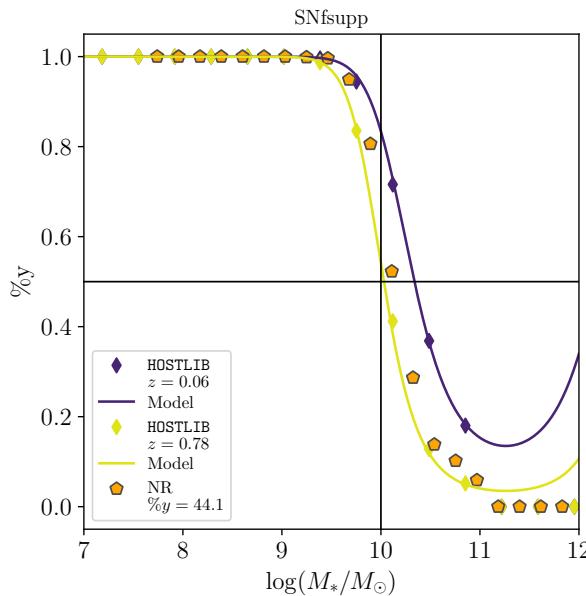
Ainsi, nous avons implémenté dans `SNANA` les différentes corrélations sous-jacentes et modélisations des propriétés des SNe Ia des études de **SK**, **BP**, **NN** et de cette thèse (NR) *via* le biais de `HOSTLIB`. Ces simulations nous permettent de générer des échantillons

corrigés des biais reproduisant les observations des sondages LOWZ, SDSS, PS1 et SNLS comprenant  $\approx 13\,000$  données. Ces différentes modélisations peuvent être combinées entre elles pour tester la qualité des hypothèses sous-jacentes et le possible biais dû au fait de mal corriger les SNe Ia. Nous traitons maintenant de la qualité d'ajustement des données simulées aux données réelles.

### I.3 Comparaison des données simulées aux données réelles

Avant de comparer les résultats cosmologiques, nous nous sommes intéressés à la correspondance entre les données simulées et les données réelles afin d'apprécier les implications sur les distributions des différentes modélisations. Nous présentons dans cette section les différents diagnostiques nous permettant de comparer à la fois graphiquement et numériquement l'accord entre les données simulées et données réelles.

En premier lieu, nous exposons Figure I.6 la fraction de jeunes étoiles en fonction de la masse pour le modèle de masse de référence, SNf<sub>supp</sub>, et pour le modèle NR.



**FIGURE I.6 – En violet (jaune) :** fraction de jeunes étoiles en fonction de la masse pour le modèle de masse SNf<sub>supp</sub> au redshift moyen de la HOSTLIB utilisée à hauts (bas) redshifts. **En orange :** même fraction mais pour l'échantillon simulé NR. Nous observons bien ici la suppression du modèle pour  $M_* > 10^{11} M_\odot$ .

Ensuite, pour avoir une comparaison efficace, nous effectuons une sélection aléatoire des données de chaque sondage pour reproduire les ratios attendus. Les quantités de données de ces mesures, exprimées en pourcentages de l'échantillon Pantheon, sont égales aux ratios du plus petit sondage simulé des données non échantillonnées : ceux de la colonne « SDSS » du Tableau I.5.

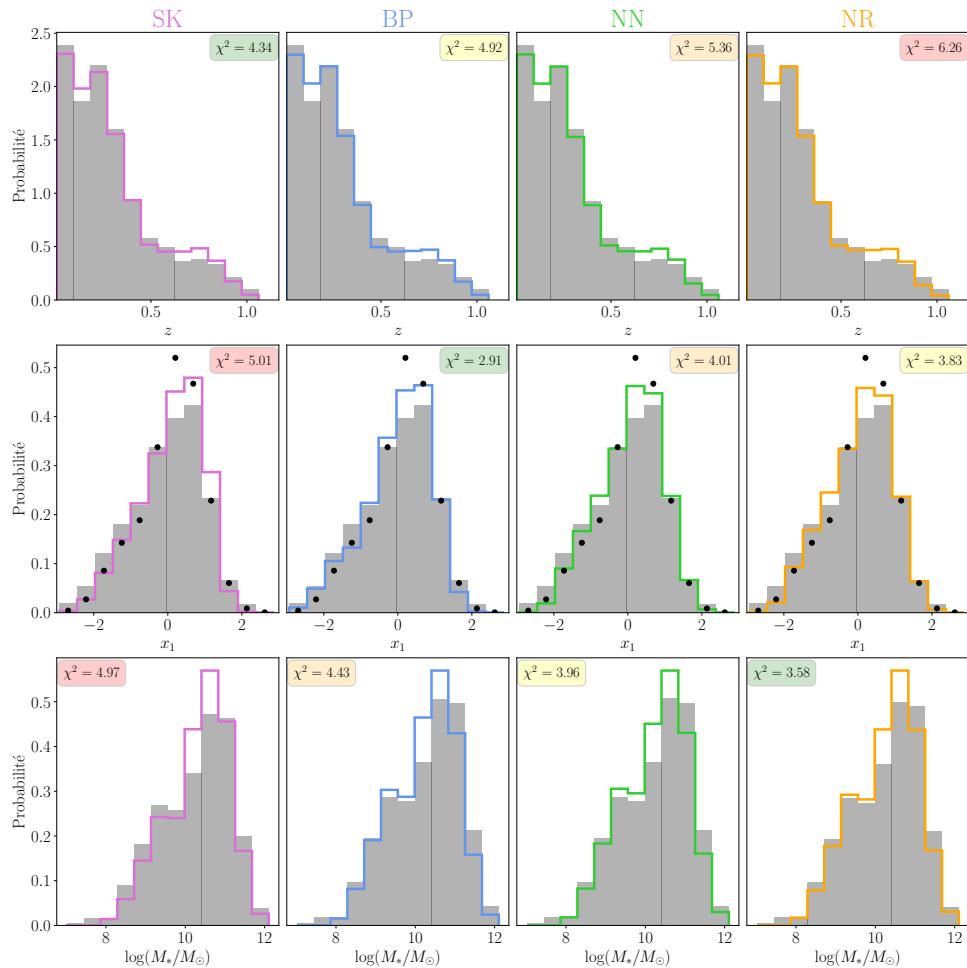
Nous rappelons que les modèles BP et SK utilisent des distributions des paramètres spécifiquement ajustés aux données ; BP utilisent des distributions gaussiennes asymétriques, avec 3 paramètres libres, dans des intervalles de  $0.2 \times 10^{10} M_\odot$  (10 pour LOWZ, 20 pour les autres) pour reproduire l'étirement des SNe Ia ; SK incluent également des distributions gaussiennes asymétriques, une pour chacun des sondages SDSS, PS1 et SNLS, et la distribution donnée Équation I.2 avec 6 paramètres libres, pour un total de  $k = 15$ . À l'inverse, les modélisations NN et NR reposent sur une modélisation prospective, basée sur une modélisation de l'étirement avec 5 paramètres libres ainsi qu'une modélisation de

la masse avec 4 paramètres libres. L'évolution de la fraction de jeunes étoiles repose sur 2 paramètres ( $K, \Phi$ ) qui sont fixés.

Nous nous intéressons dans un premier temps à l'ajustement en 1 dimension des paramètres (Section I.3.1) avant de traiter l'aspect bi-dimensionnel (Section I.3.2).

### I.3.1 Accord entre les données : analyse uni-dimensionnelle

Nous présentons ici les résultats des simulations de paramètres de redshift, étirement et masse des différentes modélisations dont les représentations graphiques sont données Figure I.7.



**FIGURE I.7** – Histogrammes normés des données simulées (en lignes pleines colorées) et des données réelles (en gris) selon le modèle et le paramètre. *De gauche à droite* : résultats pour les modèles SK, BP, NN et NR, respectivement. *De haut en bas* : nombre de données simulées en fonction du redshift, de l'étirement et de la masse, respectivement. Les valeurs de  $\chi^2$  entre les données simulées et réelles sont indiquées dans le coin supérieur droit de chaque figure, et de vert à rouge du plus petit au plus grand. Nous indiquons en points noirs le modèle d'étirement de NN au redshift moyen de l'échantillon Pantheon.

Pour chacune des comparaisons, nous calculons une valeur de  $\chi^2$ . Pour cela, nous normalisons les histogrammes des données simulées au nombre de données de Pantheon,

puis calculons :

$$\chi^2 = \frac{1}{N} \times \sum_{i=0}^{N-1} \frac{(d_i - s_i)^2}{d_i + s_i} \quad (\text{I.5})$$

avec  $N$  le nombre d'intervalles des histogrammes et  $d_i$  ( $s_i$ ) le nombre de données réelles (simulées) dans l'intervalle  $i$ . Le meilleur accord est décrit par le  $\chi^2$  le plus petit. Les valeurs sont indiquées Tableau I.6.

**Tableau I.6** – Valeurs de  $\chi^2$  donnant la comparaison de la capacité de chaque simulation à représenter les données de redshift, d'étirement et de masse.

Paramètre	$\chi^2$			
	SK	BP	NN	NR
Redshift	4.34	4.92	5.36	6.26
Étirement	5.01	2.91	4.01	3.83
Masse	4.97	4.43	3.96	3.58
Somme	14.32	12.26	13.33	13.67
Probabilité	0.36	1.00	0.59	0.49

**Notes.** Pour chaque simulation, une sélection des données est réalisée pour correspondre aux ratios des données de Pantheon, et le calcul du  $\chi^2$  est la moyenne sur 500 de ces tirages à chaque fois. La probabilité est donnée par rapport au meilleur modèle (BP), telle que  $\mathcal{P}_{\text{modèle}} = \exp(\chi^2_{\text{BP}} - \chi^2_{\text{modèle}})/2$ .

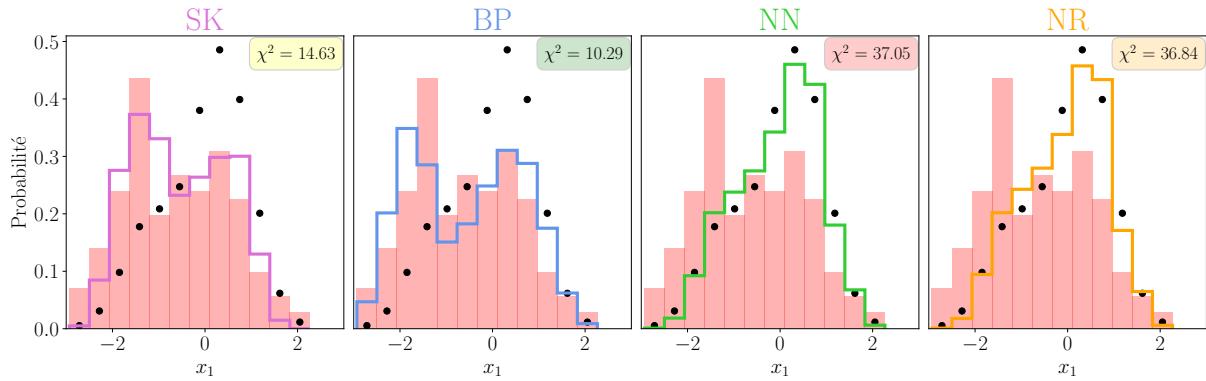
D'une manière globale, le modèle BP apparaît comme la meilleure description des données, NN et NR donnent des résultats similaires et SK a le moins bon accord. Pour le redshift cependant, c'est le modèle SK qui est le mieux ajusté aux données ; ceci correspond à nos attentes puisque leurs distributions d'étirements se divisent par redshift. Pour la masse, étant donné que toutes les simulations utilisent les mêmes WEIGHTMAP, les différences sont moins notables. NR donne cependant une meilleure représentation des données, mais pas de manière significative. Pour l'étirement, c'est BP qui décrit le mieux les données ; ceci correspond également à nos attentes puisque leurs distributions d'étirement sont nombreuses. Nous notons cependant que pour ce paramètre, les modèles NN et NR sont bien représentatifs des données. En regardant par échantillon, nous observons qu'en réalité les modèles NN et NR performent en moyenne bien mieux que les deux autres mais ont par contre une certaine difficulté à reproduire la distribution de LOWZ. En effet, par la nature ciblée du sondage, la prédiction du modèle de NN ne peut s'appliquer, ce qui mène à leurs valeurs de  $\chi^2$  que nous détaillons Tableau I.7. Nous présentons Figure I.8 l'accord entre données simulées et réelles de l'échantillon LOWZ pour les différents modèles.

Il reste que dans la pratique, tous sondages confondus, les résultats des différents modèles sont compatibles entre eux, et nous pouvons tous les considérer comme de bonnes représentations des données. Pour LOWZ spécifiquement, nous pourrions améliorer la simulation du sondage *via* la modification du modèle NN pour l'étirement, notamment en incluant les données de ZTF (Chapitre ?? Section ??) ou en utilisant la fraction de jeunes étoiles escomptée. En effet, le modèle d'évolution de la fraction de jeunes étoiles  $\delta(z)$  donne une valeur de 50% de jeunes SNe Ia à  $z = 0,05$  ; or dans notre cas, le sondage se situe à un redshift moyen de  $z = 0,03$  mais les données testées (voir Chapitre ??, Figure ??) du

**Tableau I.7** – Valeurs de  $\chi^2$  donnant la comparaison de la capacité de chaque simulation à représenter les données d'étirement pour chaque sondage simulé.

Sondage	$\chi^2$			
	SK	BP	NN	NR
LOWZ	14.63	10.29	37.05	36.84
SDSS	7.03	8.52	7.66	7.05
PS1	10.35	3.58	4.06	4.29
SNLS	15.14	23.13	15.03	15.00
Somme	47.15	45.52	63.80	63.18
Sans LOWZ	35.52	35.23	26.75	26.34

**Notes.** Ici, aucun tirage n'est réalisé.

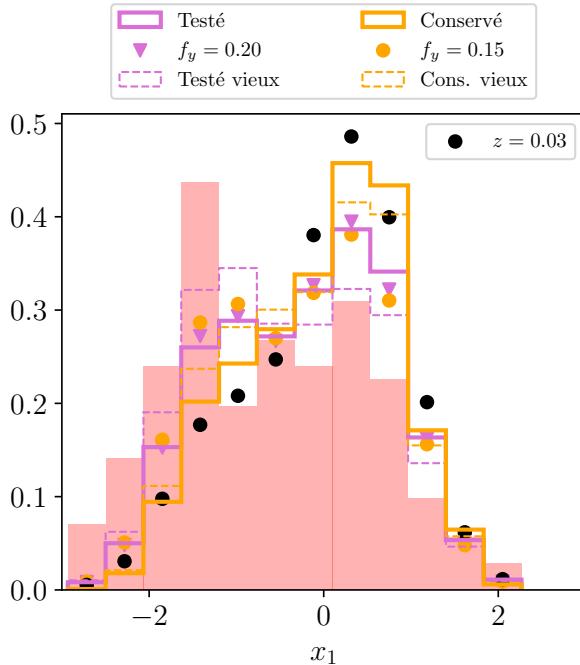


**FIGURE I.8** – Histogrammes normés des étirements des données simulées (en lignes pleines colorées) et des données réelles (en rouge) pour le sondage LOWZ selon le modèle. *De gauche à droite* : résultats pour les modèles SK, BP, NN et NR, respectivement. Les valeurs de  $\chi^2$  entre les données simulées et réelles sont indiquées dans le coin supérieur droit de chaque figure, et de vert à rouge du plus petit au plus grand. Nous indiquons en points noirs le modèle d'étirement de **NN** au redshift moyen de l'échantillon LOWZ.

modèle NR n'en possèdent que 20%, réduit à 15% dans les données conservées. Comme nous avons créé notre HOSTLIB en utilisant le redshift de chaque entrée, l'accord avec les données est de fait erroné, et il est probable que la vraie fraction soit encore plus faible en observant ces résultats Figure I.9.

### I.3.2 Accord entre les données : analyse bi-dimensionnelle

Nous présentons maintenant les distributions d'étirement en fonction du redshift d'une part et les distributions d'étirement en fonction de la masse de la galaxie hôte d'autre part. Pour déterminer l'accord entre les échantillons réels et simulés de manière quantitative, nous avons utilisé une estimation par noyau pour convertir les données simulées en densité de probabilité bi-dimensionnelle, permettant de calculer une probabilité totale traduisant l'accord entre les données réelles et le noyau. Deux exemples sont donnés Figure I.10, où nous représentons les distributions des données simulées *via* son estimation par noyau en



**FIGURE I.9** – Histogrammes des étirements des données de LOWZ : *en rouge* celles de Pantheon ; *en violet* les données testées et *en orange* les données conservées pour le modèle NR. Le modèle **NN** évalué aux fractions des jeunes SNe Ia pour ces deux échantillons sont tracés en marqueurs de la couleur correspondante ; le modèle évalué au redshift moyen de la distribution est tracé en marqueurs noirs. Les parties vieilles des données testées et conservées sont en pointillés.

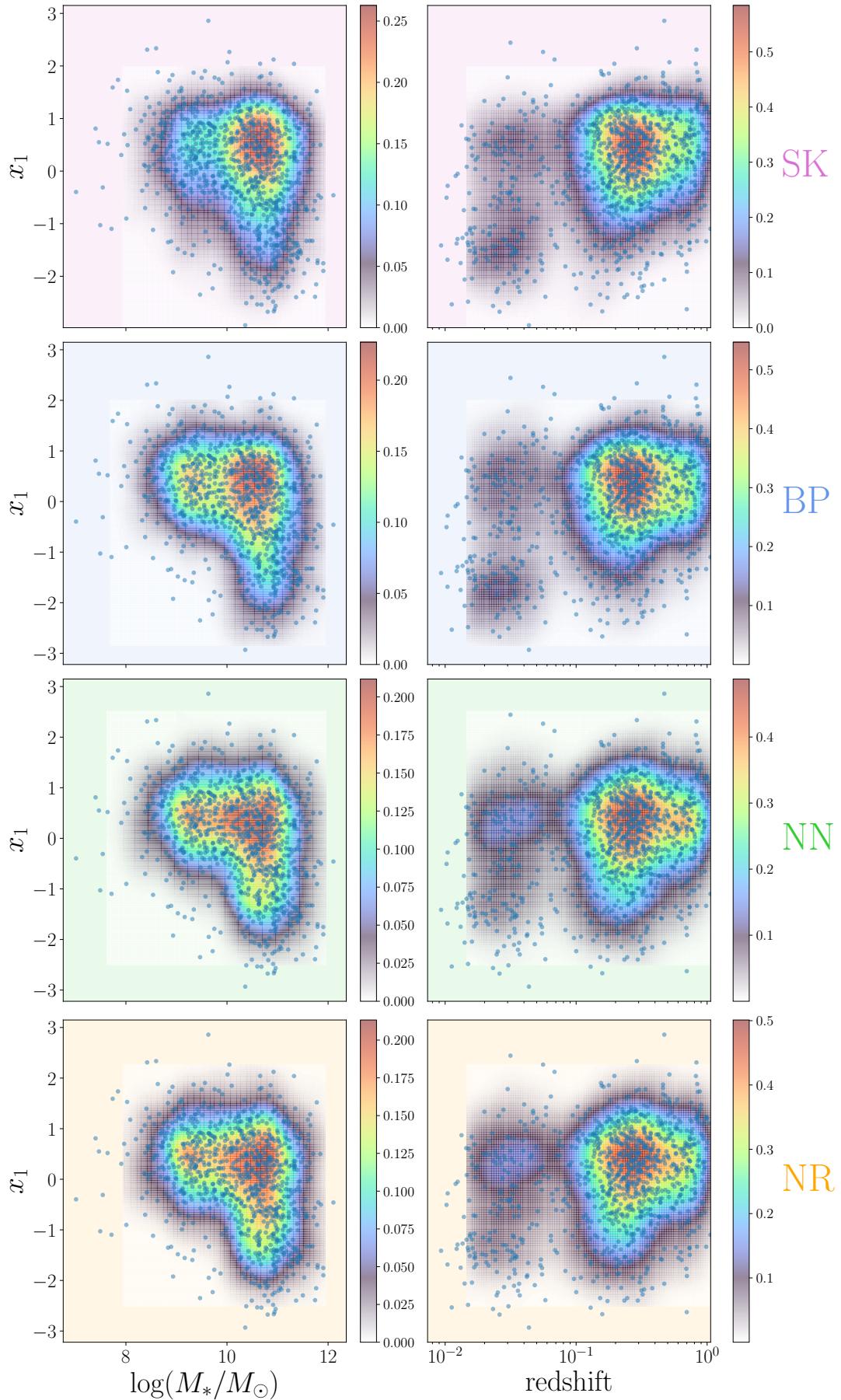
couleurs et en points dispersés pour les données réelles. Les résultats sont indiqués dans le Tableau I.8 : une plus grande probabilité représente un meilleur accord.

**Tableau I.8** – Comparaison de la capacité de chaque simulation à représenter les données d'étirement et de masse d'une part, et d'étirement et de redshift d'autre part.

Modèles	Probabilité		
	$x_1$ vs $M_*$	$x_1$ vs $z$	Somme
SK	103,03	252,57	355,60
BP	103,37	246,49	349,85
NN	102,35	236,25	338,60
NR	102,32	235,61	338,93

**Notes.** Pour chaque simulation, nous calculons une estimation par noyau bi-dimensionnelle sur les données simulées et nous l'utilisons pour déterminer chaque probabilité.

Nous observons que la modélisation **SK** est la meilleure des 4 sur la combinaison de ces distributions ; la modélisation **BP** est deuxième, et les modélisations **NN** et **NR** ont un score similaire les plaçant comme les modélisations les moins bien ajustées aux données. Les résultats ne diffèrent que très peu sur les distributions conjointes de redshift et de masse, et la majeure partie de la différence entre les modèles vient de la modélisation conjointe de l'étirement et du redshift pour laquelle les valeurs de probabilités varient plus rapidement du fait de la faible étendue des données (voir barres de couleur). Nous observons cependant que les modèles SK et BP présentent deux nuages de probabilité à peu près équivalents à  $z < 0,1$ , correspondant au sondage LOWZ, alors que pour NN et NR le nuage de point de haut étirement est plus prononcé que celui de petit étirement. Ceci découle naturellement de la différence de modélisation de l'étirement de ce sondage,



**FIGURE I.10** – Accord données réelles (en points bleus transparents) et simulées (en couleur) en 2D pour tous les modèles. *À gauche* : étirement en ordonnée et masse en abscisse. *À droite* : étirement en ordonnée et redshift en abscisse. *De haut en bas* : modèles SK, BP, NN et NR, respectivement, voir légende ; les couleurs correspondent à celles de la Figure I.7.

comme discuté précédemment. Nous notons cependant que les quatre modèles donnent un accord similaire aux données réelles.

## I.4 Impact sur la cosmologie

Maintenant que nous avons observé l'accord de chacun de modèles aux données réelles, nous pouvons étudier le biais cosmologique du fait de traiter des données ayant leur propre physique avec une correction potentiellement différente. Pour cela, comme introduit Section I.2.5, nous appliquons la méthode BBC7D ([POPOVIC et al. 2021](#), voir Chapitre ?? Section ??) sur les données des différents modèles avec chacun des échantillons de BiasCor. Nous obtenons ainsi 16 échantillons corrigés de taille  $\approx 13\,000$  (dont le nombre de données est indiqué Tableau I.5), chacun ayant une valeur de  $w$ ,  $\gamma_{\text{masse}}$ ,  $\alpha$  et  $\beta$ ; pour rappel, nous avons fixé  $\Omega_M$  à  $0,315 \pm 0,005$ , et nous ne nous intéressons donc pas à la variation de ce paramètre.

Nous présentons Section I.4.1 les résultats pour les paramètres de standardisation  $\alpha$ ,  $\beta$  et  $\gamma$ , nécessaires pour pouvoir comparer les résultats sur le paramètre d'état de l'énergie noire  $w$  dont les résultats sont présentés Section I.4.2. Nous discutons de l'impact du choix de modèle de masse Section I.4.3.

### I.4.1 Résultats de standardisation

Afin d'avoir des résultats cosmologiques significatifs, nous nous sommes intéressés aux valeurs des paramètres  $\alpha$ ,  $\beta$  et  $\gamma$  ajustées de nos simulations. Les résultats pour  $\alpha$  sont indiqués sur la Figure I.11, et ceux de  $\beta$  et  $\gamma$  sont indiqués Figure I.12; les couleurs représentent l'écart aux valeurs de référence, avec  $\gamma_{\text{ref}} = 0,05$ . Nous nous attendons à ce que les échantillons diagonaux ne présentent pas de biais de mesure, étant donné que leur correction des données est cohérente avec leur génération.

Nous soulignons que le fait que les deux dernières lignes soient les mêmes provient du fait que le procédé BBC7D ne prend pas en compte la variation de magnitude du modèle NR quand l'échantillon est utilisé comme BiasCor, puisqu'il est évalué en tant que  $\theta$  pour les données.

Afin de reproduire les résultats de l'analyse de Pantheon ([SCOLNICK et al. 2018](#)), nous voulons que les valeurs de sorties des termes diagonaux soient compatibles avec les valeurs d'entrées,  $\alpha_{\text{ref}} = 0,145$  et  $\beta_{\text{ref}} = 3,1$ . Lors de nos premières simulations, NR donnait une valeur de  $\alpha$  réduite à  $\approx 0,135$ . Nous avons dû rehausser la valeur d'entrée pour ce modèle à  $\alpha_{\text{ref},\text{NR}} = 0,155$  afin d'obtenir une valeur satisfaisante. Ce biais est montré *via* l'utilisation de  $\Delta\alpha = \alpha - \alpha_{\text{ref}}$  sur la Figure I.11b. Nous n'avons cependant pas ajusté la valeur  $\beta$  de référence pour NR, bien qu'elles soient  $\approx 2\sigma$  écartées de la valeur de référence, puisque la couleur n'est pas un paramètre que nous modifions dans la HOSTLIB NR ; celle-ci utilise en effet les paramètres de BP, et nous avançons que le biais sur  $\beta$  est dû à un effet subtile reliant la couleur à l'étirement.

Sur l'ensemble de ces figures, nous voyons que les modèles SK, BP et NN sont consistants avec eux-mêmes puisque les valeurs de sorties sont compatibles avec les valeurs d'entrée. S'il n'y a pas tant d'écart à 0,145 pour  $\alpha$  dans les termes non-diagonaux, nous notons que le modèle SK (pour lequel il n'existe pas de corrélation entre galaxie et étirement/couleur) utilisé en tant que données (colonne de gauche) présente des valeurs réduites quand il

Données					
	SK	BP	NN	NR	
BiasCor	SK	0.146 ± 0.002	0.144 ± 0.002	0.148 ± 0.002	0.148 ± 0.002
	BP	0.144 ± 0.002	0.144 ± 0.002	0.147 ± 0.002	0.146 ± 0.002
	NN	0.142 ± 0.002	0.143 ± 0.002	0.146 ± 0.002	0.145 ± 0.002
	NR	0.142 ± 0.002	0.143 ± 0.002	0.146 ± 0.002	0.145 ± 0.002
Données					
	SK	BP	NN	NR	
BiasCor	SK	0.001 ± 0.002	-0.001 ± 0.002	0.003 ± 0.002	-0.007 ± 0.002
	BP	-0.001 ± 0.002	-0.001 ± 0.002	0.002 ± 0.002	-0.009 ± 0.002
	NN	-0.003 ± 0.002	-0.002 ± 0.002	0.001 ± 0.002	-0.010 ± 0.002
	NR	-0.003 ± 0.002	-0.002 ± 0.002	0.001 ± 0.002	-0.010 ± 0.002

(a) Valeurs de  $\alpha$ .(b) Valeurs de  $\Delta\alpha$ .

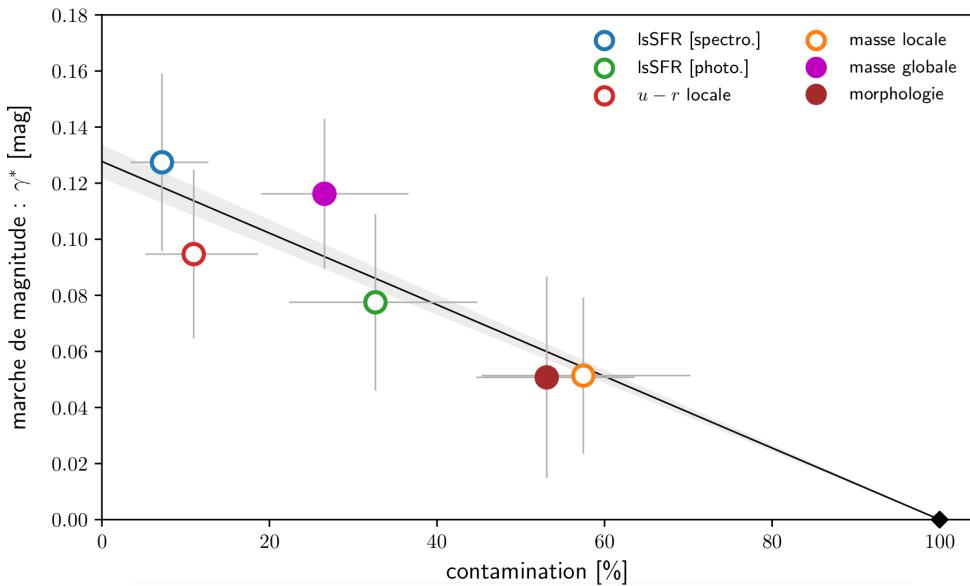
**FIGURE I.11** – Résultats cosmologiques : valeurs de  $\alpha$  et  $\Delta\alpha$  déterminées par ajustement avec la méthode BBC7D (voir Chapitre ??). La figure de droite met en évidence que le modèle NR perçoit une valeur de référence de  $\alpha = 0,155$ .

Données					
	SK	BP	NN	NR	
BiasCor	SK	3.087 ± 0.019	3.128 ± 0.019	3.137 ± 0.019	3.150 ± 0.020
	BP	3.056 ± 0.019	3.106 ± 0.019	3.120 ± 0.019	3.144 ± 0.019
	NN	3.051 ± 0.019	3.117 ± 0.019	3.122 ± 0.018	3.140 ± 0.019
	NR	3.051 ± 0.019	3.117 ± 0.019	3.122 ± 0.018	3.140 ± 0.019
Données					
	SK	BP	NN	NR	
BiasCor	SK	0.051 ± 0.003	0.036 ± 0.003	0.038 ± 0.003	0.094 ± 0.003
	BP	0.061 ± 0.003	0.044 ± 0.003	0.048 ± 0.003	0.103 ± 0.003
	NN	0.060 ± 0.003	0.045 ± 0.003	0.049 ± 0.003	0.104 ± 0.003
	NR	0.060 ± 0.003	0.045 ± 0.003	0.049 ± 0.003	0.104 ± 0.003

(a) Valeurs de  $\beta$ .(b) Valeurs de  $\gamma$ .

**FIGURE I.12** – Résultats cosmologiques : valeurs de  $\beta$  et  $\gamma$  déterminées par ajustement avec la méthode BBC7D (voir Chapitre ??).

est corrigé par les autres modèles (qui, eux, supposent des corrélations) ; celles-ci sont augmentées quand SK est utilisé comme BiasCor (ligne du haut). Ce même résultat se retrouve dans pour les valeurs de  $\beta$  (Figure I.12a), mais ce phénomène est inversé pour les valeurs de  $\gamma$  (Figure I.12b) : la colonne de gauche présente des valeurs plus hautes et la ligne du haute des valeurs plus basses. Ces effets sont corrélés puisque l'absence de



**FIGURE I.13** – Marche de magnitude des SNe Ia en fonction du traceur. Figure de [BRIDAY \(2021\)](#). Les valeurs des marches de magnitude (points ouverts pour les traceurs locaux et pleins pour les globaux) sont déterminées sur l'échantillon SNf. Selon la capacité des traceurs à distinguer les populations de SNe Ia (traduit par leur contamination), la marche de magnitude trouvée varie linéairement suivant la droite noire (avec son erreur en bande grise). En prenant le LsSFR comme traceur de référence avec  $\gamma = 0,13$  mag, la masse globale donnerait une marche de magnitude autour de 0,10 mag sur cette ligne.

corrélation entre galaxie et étirement/couleur du modèle SK est compensé dans la valeur de marche de magnitude.

Pour NR, nous supposons que le fait que nous devons modifier la valeur de  $\alpha_{\text{ref}}$  provient d'une mauvaise interprétation de la corrélation de la magnitude avec l'âge par **SNANA**; le programme corrige cette magnitude *via* la valeur de la masse de la galaxie hôte, et ne possède pas toute l'information nécessaire pour comprendre cette corrélation. Ceci se répercute sur les résultats de  $\gamma_{\text{masse}}$ ; en effet, alors que sa valeur d'entrée de marche de magnitude basée sur l'âge est de 0,130 mag, l'échantillon NR en tant que données (colonne de droite Figure I.12b) donne des valeurs de marche de magnitude basées sur la masse ajustées par la méthode BBC7D de  $\gamma_{\text{masse}} \approx 0,100$ , soit deux fois plus élevée que les valeurs trouvées pour les autres simulations (de valeur d'entrée 0,05 mag).

Cette déviation est attendue étant donné que la masse constitue un mauvais traceur de l'âge d'une SN. Il est cependant intéressant de voir ces valeurs plus basses que celle d'entrée : c'est d'une part cohérent avec le fait que la simulation voie un  $\alpha_{\text{ref}} = 0,155$  (la baisse de  $\gamma$  est compensée dans  $\alpha$ , comme pour SK en tant que BiasCor), et d'autre part ce résultat est en total accord avec l'étude de [BRIDAY et al. \(2022\)](#). En effet, l'auteur rapporte dans cet article qu'un traceur moins efficace à différencier les populations de SNe Ia attribuera une marche de magnitude plus faible qu'un traceur plus discriminant; notamment, dans leur figure 4 (recopiée Figure I.13), une valeur de  $\gamma = 0,13$  mag avec le LsSFR se traduirait par une valeur de  $\gamma \approx 0,10$  mag avec la masse globale (sur la ligne noire sous la masse globale).

### I.4.2 Résultats de cosmologie

Dans la section précédente, nous nous sommes intéressés aux corrélations intrinsèques aux SNe Ia impactant les paramètres de standardisation. Dans cette section, nous présentons les résultats des valeurs de  $w$  ajustées par `wfit` (voir Section ??) avec une valeur fixée de  $\Omega_M = 0,315$ ; dans ce cas, ce sont des évolutions avec le redshift (erronées ou non) qui impactent la valeur de  $w$ . Ainsi, nous présentons Figure I.14 les résultats des valeurs ajustées de  $w$ , où les couleurs représentent l'écart à la valeur de référence ( $w = -1,00$ ), et pour comprendre ces résultats, nous avons observé les évolutions des résidus de HUBBLE (voir Chapitre ??) en fonction du redshift pour chaque échantillon, représentées Figure I.15. De même que précédemment, nous nous attendons à ce que les termes diagonaux soient compatibles avec  $-1,00$ , et à ce que les deux dernières lignes des figures soient similaires puisque les modèles NN et NR forment les mêmes BiasCor.

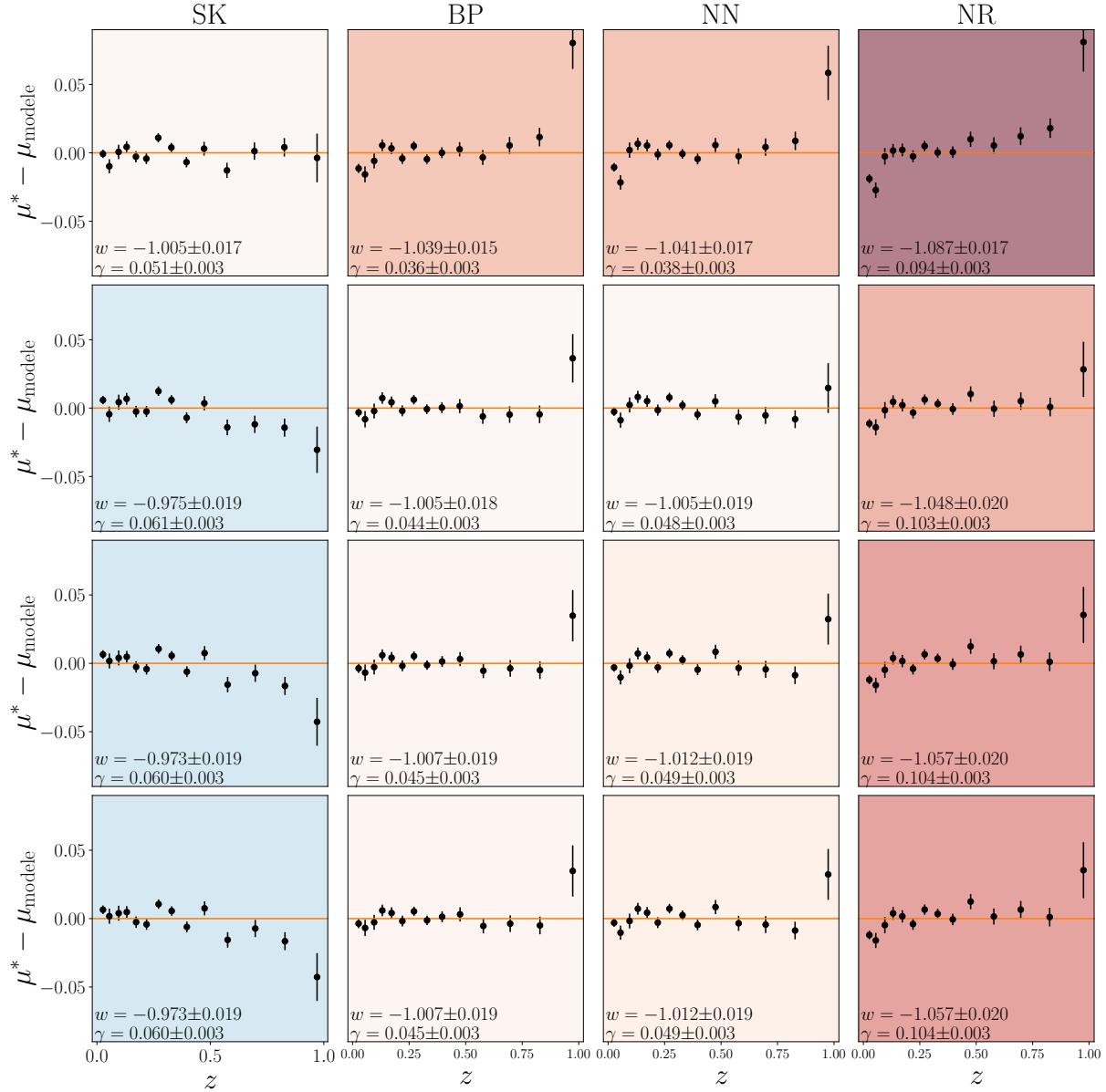
		Données			
		SK	BP	NN	NR
BiasCor	SK	-1.005 ± 0.017	-1.039 ± 0.015	-1.041 ± 0.017	-1.087 ± 0.017
	BP	-0.975 ± 0.019	-1.005 ± 0.018	-1.005 ± 0.019	-1.048 ± 0.020
	NN	-0.973 ± 0.019	-1.007 ± 0.019	-1.012 ± 0.019	-1.057 ± 0.020
	NR	-0.973 ± 0.019	-1.007 ± 0.019	-1.012 ± 0.019	-1.057 ± 0.020

**FIGURE I.14** – Valeurs de  $w$  déterminées par ajustement avec la méthode BBC7D (voir Chapitre ??).

Pour rappel, nous définissons  $w$ , le paramètre d'état de l'énergie sombre, *via* l'équation mathématique reliant la pression du fluide parfait la décrivant au facteur d'échelle  $a$  de l'Univers :

$$\rho(t) \propto a(t)^{-3(1+w)} \quad (\text{I.6})$$

pour laquelle  $w = -1$  donne une valeur de la pression constante dans le temps, alors qu'une valeur de  $w$  plus grande que  $-1,00$  (par exemple,  $-0,95$ ) impliquerait que l'effet de l'énergie sombre diminue avec le temps ( $\rho(t)$ ) de la forme  $\rho(t) \propto a(t)^{-x}$  avec  $x > 0$ ) et une valeur plus petite en augmenterait la puissance ( $\rho(t) \propto a(t)^{+x}$  avec  $x > 0$ ); autrement dit, si  $w > -1$  alors l'énergie sombre était plus puissante dans le passé et le sera moins dans le futur, et donc que l'Univers est en réalité plus jeune que ce que l'on croit, et inversement. Les données actuelles de Pantheon (SCOLNICK et al. 2018) contraint  $w$  à  $\pm 0,220$  à elles seules, et à  $\pm 0,040$  avec d'autres sondes, donnant des résultats compatibles avec  $-1,00$ ; les futurs relevés visent une détermination à  $\pm 0,020$ , ce qui pourrait mettre en évidence un



**FIGURE I.15** – Résultats cosmologiques : résidus de HUBBLE déterminés par ajustement avec la méthode BBC7D (voir Chapitre ??).

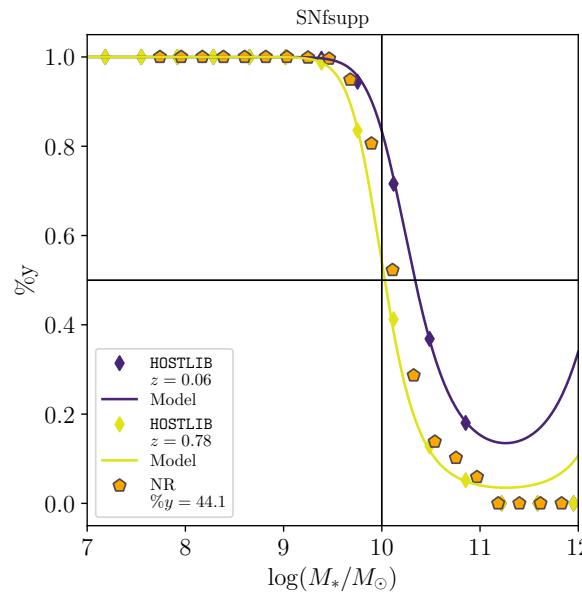
décalage à  $w = -1,00$ . À cet effet, nos simulations sont représentatives de ce qui pourra être atteint dans un futur proche.

Sur la première figure, nous observons que les trois premiers modèles corrigés de manière cohérente avec leur génération (sur la diagonale donc) sont tous compatibles avec  $w = -1,00$ , bien que NN s'en écarte un peu. En revanche, le modèle SK rend incompatible les autres modèles, que ce soit en l'utilisant pour les données (colonne de gauche) ou en tant que BiasCor (ligne du haut), introduisant un biais jusqu'à 4% pour l'échantillon NN\_SK. Si cela ressemble aux résultats de standardisation, ici cela résulte d'une sous-ou sur-correction systématique selon l'utilisation de SK, comme nous pouvons le voir Figure I.15 : les modèles donnant  $w = -1,00$  sont équilibrés autour de 0, alors qu'une sous-correction augmente cette valeur. Les échantillons BP\_NN et NN\_BP, en revanche, donnent des valeurs compatibles avec  $-1,00$ ; étant donné que les résultats cosmologiques actuels se basent sur un approche semblable à celle de ?, ce résultat conforte la validité du modèle NN.

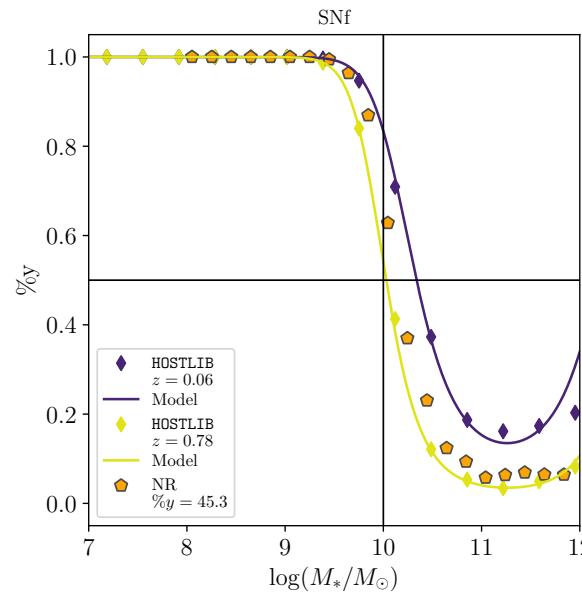
Alors que les trois premiers modèles utilisent la masse comme traceur de la marche de magnitude et ne contiennent pas d'évolution du vrai  $\gamma$  avec le redshift, le modèle NR utilise l'âge et donc présente une évolution de  $\gamma$  avec le redshift. En effet, à  $z = 0,05$ , il est attendu 50% de jeunes étoiles; ainsi à ce redshift, deux SNe Ia auront soit aucune déviation due à une marche de magnitude (deux jeunes ou deux vieilles), soit une différence de magnitude entre elles de 0,130 mag, soit une moyenne à 0,065 mag; à l'inverse, à haut redshift les SNe Ia sont principalement jeunes, donc la différence de magnitude entre 2 SNe Ia tend en moyenne vers 0 mag. Par rapport à la marche de magnitude basée sur la masse, ce modèle sur-évalue la marche à bas redshift et la sous-évalue à haut redshift. Ceci est visible sur la Figure I.15 où, dans la colonne de droite, les points à bas redshift sont plus bas que dans les autres modèles et les points à haut redshift ( $z > 0,50$ ) sont réhaussés. Ceci explique le fait que le modèle NR utilisé comme données ne fournit pour sa part aucun résultat de  $w$  proche de  $-1,00$  dans la Figure I.14 : SNANA ne possède pas tous les outils pour prendre en compte toutes les implications du fait d'utiliser l'âge comme traceur des propriétés des SNe Ia. Dans ces conditions, nous observons un biais sur la mesure de  $w$  aux alentours de 5% si les données présentent effectivement des corrélations avec l'âge d'une SN mais que la correction utilisée ne le suppose pas (résultat NR\_BP par exemple), et jusqu'à 8% si la correction n'inclut aucune corrélation à la galaxie (résultat NR\_SK).

### I.4.3 Systématiques dues au choix du modèle de masse

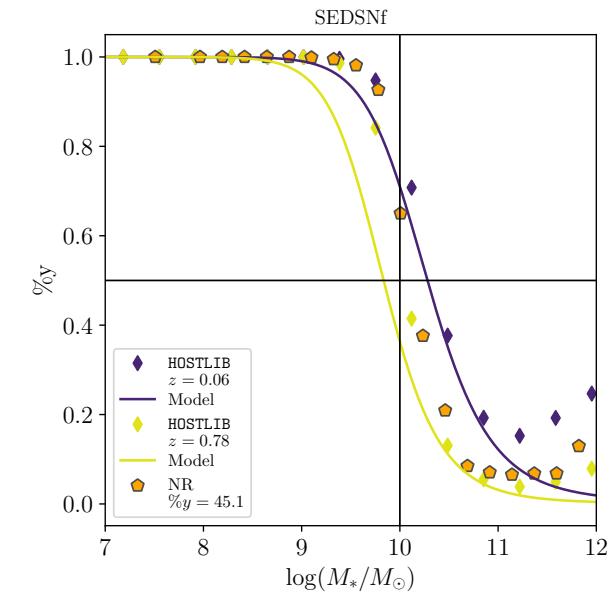
Le choix du modèle de masse diffère principalement dans la fraction retrouvée de jeunes étoiles en fonction de la masse, dont nous donnons les représentations graphiques Figure I.16. Nous exposons ici les différences cosmologiques dues au choix de la modélisation de la masse, présentées sur les valeurs de  $w$  Figure I.17.



(a) Fraction pour SNfsupp.



(b) Fraction pour SNf.



(c) Fraction pour SEDSNf.

**FIGURE I.16** – De gauche à droite : modèles SNfsupp, SNf et SEDSNf. En violet (jaune) : fraction de jeunes étoiles en fonction de la masse au redshift moyen de la HOSTLIB utilisée à hauts (bas) redshifts ; en orange : même fraction mais pour l'échantillon simulé NR.

Données				
	SK	BP	NN	NR
SK	-1.005 ± 0.017	-1.039 ± 0.015	-1.041 ± 0.017	-1.087 ± 0.017
BP	-0.975 ± 0.019	-1.005 ± 0.018	-1.005 ± 0.019	-1.048 ± 0.020
NN	-0.973 ± 0.019	-1.007 ± 0.019	-1.012 ± 0.019	-1.057 ± 0.020
NR	-0.973 ± 0.019	-1.007 ± 0.019	-1.012 ± 0.019	-1.057 ± 0.020

(a) Valeurs de  $w$  pour SNf.

Données				
	SK	BP	NN	NR
SK	-1.005 ± 0.017	-1.039 ± 0.015	-1.037 ± 0.015	-1.074 ± 0.015
BP	-0.975 ± 0.019	-1.005 ± 0.018	-1.004 ± 0.018	-1.037 ± 0.018
NN	-0.980 ± 0.019	-1.010 ± 0.019	-1.005 ± 0.018	-1.042 ± 0.018
NR	-0.980 ± 0.019	-1.010 ± 0.019	-1.005 ± 0.018	-1.042 ± 0.018

(b) Valeurs de  $w$  pour SNf.

Données				
	SK	BP	NN	NR
SK	-1.005 ± 0.017	-1.039 ± 0.015	-1.026 ± 0.015	-1.078 ± 0.015
BP	-0.975 ± 0.019	-1.005 ± 0.018	-0.991 ± 0.016	-1.039 ± 0.016
NN	-0.973 ± 0.019	-1.010 ± 0.018	-0.997 ± 0.018	-1.048 ± 0.018
NR	-0.973 ± 0.019	-1.010 ± 0.018	-0.997 ± 0.018	-1.048 ± 0.018

(c) Valeurs de  $w$  pour SEDSNf.

**FIGURE I.17** – Résultats cosmologiques : valeurs de  $w$  pour les modèles de masse SNfsupp à gauche, SNf au milieu et SEDSNf à droite.

Nous observons que les différents modèles de masse donnent des fractions de jeunes étoiles totale similaires entre eux ( $\approx 45\%$ ) ; leur distribution n'est cependant pas la même. En effet, à  $M_* > 10^{11} M_\odot$  le modèle SNfSupp est (par construction) dénué de jeunes étoiles, mais les WEIGHTMAP ne favorisent pas ces masses-là donc la différence n'est pas si notable. En revanche, à  $M_* \approx 10^{10} M_\odot$ , partie privilégiée par les WEIGHTMAP, le modèle SEDSNf a la plus haute fraction de jeunes étoiles (point pentagonal orange Figure I.16).

Cette variation pourrait influencer les valeurs de  $\alpha$  et  $\beta$ , mais ceci n'est pas présent de manière notable. En revanche, elle semble se refléter dans les valeurs de  $w$  comme l'indique la Figure I.17, où nous y observons une augmentation moyenne des valeurs dans les deux colonnes de droites. Les observations précédentes restent cependant valables : le modèle NN est cohérent avec lui-même et avec le modèle BP, les corrélations entre l'âge d'une SN et ses propriétés ne sont pas traitées optimalement par SNANA et le biais potentiel sur la mesure de  $w$  se trouverait entre 4 et 8%. Nous pouvons estimer que le choix du modèle de masse contribue à  $\approx 0,01$  de la dispersion de  $w$ , et que les effets de l'âge sur la marche de magnitude et l'évolution avec le redshift sont les effets dominants.

## I.5 Conclusion

Nous avons présenté la suite de notre étude de la dérive de la distribution d'étirement sous-jacente des SNe Ia en fonction du redshift. Nous avons utilisé des échantillons de données de masse afin d'établir un modèle de dérive de masse qui dépend de la fraction attendue des SNe Ia jeunes et vieilles en fonction du redshift. Celui-ci nous a permis d'augmenter les HOSTLIB de BP en associant un âge à chacune des entrées, et ce faisant y changer la corrélation avec l'étirement et la magnitude.

Cette implémentation nous a permis de tester les résultats cosmologiques trouvés par SNANA lorsque la correction des données peut être faite de manière incohérente avec sa génération, en reproduisant les corrélations de SK et BP. Nous obtenons alors 16 échantillons pour lesquels nous avons des valeurs de  $\alpha$ ,  $\beta$ ,  $\gamma_{\text{masse}}$  et  $w$ .

À travers cette analyse et l'implémentation du modèle NN d'un bout à l'autre du *pipeline* d'analyse cosmologique SNANA, nous avons pu mettre en lumière les éléments suivants :

- 1) Nous avons testé la robustesse du modèle d'évolution de l'étirement avec le redshift développé dans NICOLAS et al. (2021) et présenté Chapitre ??, qui a été utilisé avec succès pour reproduire les sondages totaux et non coupés en redshift ; celui-ci s'est notamment avéré être une meilleure description des données à haut redshift que les modèles SK ou BP, confortant les indices en faveur de cette évolution ;
- 2) chacune des modélisations (SK, BP, NN, NR) sont de bonnes représentations des données. La nature ciblée du sondage LOWZ en fait un échantillon que nous n'avons pas pu reproduire dans l'état avec notre approche de notre modèle prospectif d'étirement mais il pourrait y être ajusté ;
- 3) nous avons mis en évidence la limite au fait de considérer la masse comme un traceur efficace des propriétés dérivant de l'âge d'une SN ;
- 4) nous trouvons que  $\gamma_{\text{masse}}$  est réduit par rapport à  $\gamma_{\text{âge}}$ , de manière consistante avec BRIDAY et al. (2022) ;

- 5) nous avons montré que les modèles simulés et corrigés de la même manière ne présentent pas de biais dans la récupération des valeurs de  $w$ , mais qu'il existe un biais aux alentours de 5% dans l'hypothèse d'une modélisation erronée (NR\_NR) ;
- 6) par rapport à l'état de l'art actuel des analyses cosmologiques (BP), il pourrait y avoir un biais de l'ordre de 4% si l'âge constitue le paramètre à l'origine des propriétés des SNe Ia. Cette valeur, compatible avec les incertitudes actuelles, pourra se révéler critique à l'ère des futurs grands relevés cosmologiques comme LSST qui apporteront rapidement  $\approx 15\,000$  données cosmologiques de SNe Ia.

Ces biais pourraient être corrigés avec l'implémentation de l'âge comme paramètre décrivant la physique intrinsèque des SNe Ia, et ouvrirait la voie à l'étude du biais cosmologique issu de l'utilisation de différents traceurs.

## Figures

---

I.1 $M_*$ en fonction du LsSFR des SNe Ia de SNfactory et modèle de masse sélectionné ajusté . . . . .	5
I.2 $\Delta\text{AIC}$ entre le modèle Howell et les autres modèles . . . . .	7
I.3 Modèles implémentés et testés dans l'étude de l'évolution de l'étirement avec le redshift . . . . .	8
I.4 Comparaison de la prédiction de l'évolution de la fraction de jeunes SNe Ia en fonction de la masse de la galaxie hôte . . . . .	9
I.5 Représentation du choix de l'âge d'une SN et de l'assignation de masse et d'étirement en fonction du redshift . . . . .	10
I.6 Fraction de jeunes étoiles en fonction de la masse pour le modèle de masse SNfsupp . . . . .	12
I.7 Histogrammes uni-dimensionnels des données simulées et réelles . . . . .	13
I.8 Histogrammes uni-dimensionnels des étirements des données simulées et réelles pour l'échantillon LOWZ . . . . .	15
I.9 Histogrammes des données testées et conservées du modèle NR pour le sondage LOWZ . . . . .	16
I.10 Accord entre les données réelles et simulées en 2 dimensions pour tous les modèles . . . . .	17
I.11 Résultats cosmologiques : $\alpha$ et $\Delta\alpha$ . . . . .	19
I.12 Résultats cosmologiques : $\beta$ et $\gamma$ . . . . .	19
I.13 Marche de magnitude des SNe Ia en fonction du traceur . . . . .	20
I.14 Résultats cosmologiques : $w$ . . . . .	21
I.15 Résultats cosmologiques : résidus de HUBBLE . . . . .	22
I.16 Évolution de la fraction de jeunes étoiles en fonction de la masse pour les différents modèles de masse . . . . .	24
I.17 Résultats cosmologiques : $w$ selon le modèle de masse . . . . .	24

---

## Tableaux

---

I.1 Paramètres des distributions d'étirement et de couleur pour les simulations SK . . . . .	2
I.2 Paramètres de la distribution d'étirement pour l'échantillon LOWZ des simulations SK . . . . .	3
I.3 Comparaison de la capacité relative de chaque modèle à décrire les données selon l'échantillon d'ajustement . . . . .	6
I.4 Valeurs des paramètres issus des meilleurs ajustement du modèle Howell sur les échantillons SNf et SEDSNf . . . . .	9
I.5 Nombre de données de nos différentes simulations . . . . .	11
I.6 Comparaison de la capacité de chaque simulation à représenter les données en une dimension . . . . .	14
I.7 Comparaison de la capacité de chaque simulation à représenter les données d'étirement selon le sondage . . . . .	15
I.8 Comparaison de la capacité de chaque simulation à représenter les données en deux dimensions . . . . .	16

---



## Bibliographie

- BRIDAY M. 2021, « Étude de l'impact de l'environnement galactique sur la standardisation des Supernovae de Type Ia », Theses, Université Claude Bernard – Lyon I, [HAL thèses](#)  
↑ Section I.13
- BRIDAY M., RIGAULT M., GRAZIANI R. et al. 2022, « Accuracy of environmental tracers and consequences for determining the Type Ia supernova magnitude step », [A&A](#), [657, A22](#) ↑ Section I.2, ↑ Section I.2.3, ↑ Section I.4.1, ↑ Section 4)
- BURNHAM, KENNETH & ANDERSON D. R. 2004, « Multimodel Inference: Understanding AIC and BIC in Model Selection », [Sociological Methods & Research](#), [33, 261](#) ↑ Section I.2.2
- HOWELL D. A., SULLIVAN M., CONLEY A. et CARLBERG R. 2007, « Predicted and Observed Evolution in the Mean Properties of Type Ia Supernovae with Redshift », [ApJ](#), [667, L37](#) ↑ Section I.2.1
- NICOLAS N., RIGAULT M., COPIN Y. et al. 2021, « Redshift evolution of the underlying type Ia supernova stretch distribution », [A&A](#), [649, A74](#) ↑ Section I.1.3, ↑ Section I.2.2, ↑ Section I.2.5, ↑ Section I.7, ↑ Section I.3.1, ↑ Section I.8, ↑ Section I.9, ↑ Section I.3.2, ↑ Section I.5, ↑ Section 1)
- POPOVIC B., BROUT D., KESSLER R., SCOLNIC D. et LU L. 2021, « Improved Treatment of Host-galaxy Correlations in Cosmological Analyses with Type Ia Supernovae », [ApJ](#), [913, 49](#) ↑ Section I.1, ↑ Section I.1.2, ↑ Section I.2.4, ↑ Section I.2.5, ↑ Section I.3, ↑ Section I.3.2, ↑ Section I.4, ↑ Section I.5, ↑ Section 6)
- 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](#) ↑ Section I.1.4, ↑ Section I.2.1, ↑ Section I.1
- SCOLNIC D. et KESSLER R. 2016, « Measuring Type Ia Supernova Populations of Stretch and Color and Predicting Distance Biases », [ApJ](#), [822, L35](#) ↑ Section I.1, ↑ Section I.1.1, ↑ Section ??, ↑ Section I.2, ↑ Section I.2.5, ↑ Section I.2.5, ↑ Section I.3, ↑ Section I.3.2, ↑ Section I.5
- 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](#) ↑ Section ??, ↑ Section I.1.1, ↑ Section ??, ↑ Section I.5, ↑ Section I.5, ↑ Section I.4.1, ↑ Section I.4.2
- TAYLOR E. N., HOPKINS A. M., BALDRY I. K. et al. 2011, « Galaxy And Mass Assembly (GAMA): stellar mass estimates », [MNRAS](#), [418, 1587](#) ↑ Section I.2.1