



IRAP (Research Institute in Astrophysics and Planetology)

Forecasting on next wide cosmological  
surveys of Stage IV : DESI - LSST - EUCLID

DOURNAC Fabien

GAHEC : Galaxies, High Energy Astrophysics and Cosmology



---

# Résumé

## Performances des futurs grands relevés cosmologiques : Surveys DESI et LSST

La collaboration internationale impliquée dans l'étude et la compréhension de l'expansion de l'Univers va prochainement bénéficier des données récoltées par les futurs grands surveys que sont les missions EUCLID en tant que télescope spatial et les télescopes au sol comme DESI (Dark Energy Spectroscopic Instrument) et LSST (The Large Synoptic Survey Telescope). Cette prochaine génération de surveys fait partie de l'étape dite "Stage IV" DES (Dark Energy Survey) dont les données seront collectées à peu près jusqu'en 2030 et même au-delà.

L'objectif est de reconstruire l'histoire de cette expansion et de mieux connaître l'origine de son accélération. Ceci passe par le modèle standard  $\Lambda$ CDM qui introduit une énergie noire (Dark Energy) avec une équation d'état reliant la pression et la densité d'énergie par un facteur  $w = -1$ . Cette pression de valeur négative pourrait expliquer l'augmentation de la constante de Hubble au fil du temps. Cependant, des zones d'ombre subsistent concernant sa dépendance en redshift.

Pour anticiper la mise en opération de ces différents surveys, nous disposons d'outils théoriques et numériques, en particulier le formalisme de Fisher, qui nous permettront d'évaluer leurs performances et prédire, dans une mesure relative, les contraintes sur les paramètres cosmologiques qui seront obtenues. Cette chaîne de traitement s'appelle le "Forecasting", dont la traduction française "Prévision/Prédiction" prend toute sa signification car nous "prévoyons" les incertitudes sur les mesures des grandeurs physiques en fonction des spécifications de chaque instrument.

Les deux Forecasting étudiés ici concernent les futurs relevés des télescopes DESI et LSST. Les résultats obtenus nous permettront d'envisager, selon les différents paramètres pris en compte, l'efficacité de chaque survey, autant sur la distribution 3D des galaxies (Galaxy Clustering) que sur leur déformation apparente (Weak Lensing). Le télescope DESI devrait être pleinement fonctionnel courant 2020 et la première lumière pour LSST est prévue pour la même année avant un début de collecte espéré pour 2022.



---

# Abstract

## Forecasting on next wide cosmological surveys : DESI and LSST

The international community involved in the study and understanding of the expansion of the Universe will soon benefit from the data collected by future major surveys that are the EUCLID missions as a space telescope and ground-based telescopes like DESI (Dark Energy Spectroscopic Instrument) and LSST (The Large Synoptic Survey Telescope). This next generation of surveys makes part of era called "Stage IV" DES (Dark Energy Survey) whose data will be gathered roughly up to 2030 and even beyond.

The objective is to reconstruct the history of this expansion and to better know the origin of its acceleration. This goes through the standard model  $\Lambda$ CDM that introduces a "Dark Energy" with a state equation implying a factor  $w = -1$  between pressure and energy density. Its negative value pressure could explain increasing Hubble's constant over time but some doubts remain about its redshift dependence.

To anticipate the future results of these different surveys, we have theoretical and numerical tools, in particular Fisher's formalism, which will allow us to evaluate their performances and predict, to a relative extent, the constraints on the cosmological parameters that will be got. This processing chain is called a "Forecasting" and depends on the specifications of each instrument that will collect the data.

The two Forecasting studied here concern the future surveys of the DESI and LSST telescopes. The obtained results will provide, according to the different parameters taken into account, the efficiency of each survey, as well the 3D distribution of galaxies (Galaxy Clustering) as their visible deformation (Weak Lensing). The DESI telescope should be fully functional currently in 2020 while first light for LSST will occur in the same year with a start of data collection expected for 2022.

# Table des matières

<b>Introduction</b>	<b>6</b>
<b>1 Modèle standard <math>\Lambda</math>CDM</b>	<b>7</b>
1.1 Équations de Friedmann-Lemaître	7
1.2 Énergie noire dans le modèle classique	7
1.3 Histoire de l'expansion de l'Univers	7
1.4 Extension de l'équation d'état de l'Énergie noire	9
<b>2 Forecasting</b>	<b>10</b>
2.1 Principe	10
2.2 Formalisme de Fisher	10
2.3 Critère de performances : Figure of Merit	11
2.4 Code de Boltzmann	12
<b>3 Forecasting sur le survey DESI</b>	<b>13</b>
3.1 Spécifications	13
3.2 Galaxy clustering spectroscopique	13
3.3 Paramètres cosmologiques finaux à contraindre	14
3.4 Projection	14
3.5 Jeu de paramètres choisi pour la construction de la matrice de Fisher	14
3.6 Expression du spectre de puissance observable	15
3.7 Cas population unique et multi-populations	16
3.7.1 Construction matrice de Fisher avec population unique	16
3.7.2 Construction matrice de Fisher avec multiples populations	16
3.7.3 Cas Univers plat et non plat	17
3.7.4 Résultats unique-multi populations	17
3.7.4.1 Méthode population unique - fusion de colonnes	17
3.7.4.2 Méthode population unique - moyenne pondérée	19
3.7.4.3 Interprétation des méthodes "single/multi populations"	19
3.8 Codage Python population unique et cross-corrélations	20
3.8.1 Implémentation population unique	20
3.8.1.1 Implémentation "single population" : fusion de colonnes	21
3.8.1.2 Implémentation "single population" : moyenne pondérée	21
3.8.2 Implémentation cross-corrélations	22
3.8.3 Optimisation GPU	23
3.8.4 Tâches principales et validation de l'implémentation	23
<b>4 Forecasting sur le survey LSST</b>	<b>24</b>
4.1 Spécifications	24
4.2 Weak Lensing	24
4.2.1 Caractéristiques du phénomène de Weak Lensing	24
4.2.2 Formalisme de Fisher appliqué au Weak Lensing	27
4.3 Formalisme de Fisher appliqué au Galaxy clustering photométrique	28

4.4	Formalisme de Fisher appliqué à la combinaison du Weak Lensing avec le Galaxy Clustering photométrique . . . . .	28
4.5	Synthèse finale des FoM de LSST . . . . .	28
4.6	Synthèse finale des matrices de Fisher . . . . .	29
4.7	Comparaison des résultats pour LSST . . . . .	29
4.8	Synthèse des résultats DESI-LSST . . . . .	30
4.9	Comparaison avec le survey EUCLID . . . . .	31
<b>5</b>	<b>Conclusion</b>	<b>33</b>
	<b>Bibliographie</b>	<b>35</b>
	<b>Appendice</b>	
<b>A</b>	<b>Galaxy clustering - Démonstration de l'expression du spectre de puissance observable</b>	<b>36</b>
<b>B</b>	<b>Formalisme de Fisher appliqué au Weak Lensing</b>	<b>39</b>

# Liste des Figures

1.1	<i>Facteur d'échelle normalisé de l'Univers en fonction du temps cosmique (1 Gyr = 1 Milliard d'années)</i> . . . . .	8
3.1	<i>Contours représentant les distributions posterior avec population unique "fusion de colonnes" et multi dans le cas flat (<math>k_{\max} = 0.25 \text{ h.Mpc}^{-1}</math>)</i> . . . . .	17
3.2	<i>Contours représentant les distributions posterior avec population unique "fusion de colonnes" et multi dans le cas non-flat (<math>k_{\max} = 0.25 \text{ h.Mpc}^{-1}</math>)</i> . . . . .	18
3.3	<i>Contours représentant les distributions posterior avec population unique "moyenne pondérée" et multi dans le cas flat (<math>k_{\max} = 0.25 \text{ h.Mpc}^{-1}</math>)</i> . . . . .	19
3.4	<i>Contours représentant les distributions posterior avec population unique "moyenne pondérée" et multi dans le cas non-flat (<math>k_{\max} = 0.25 \text{ h.Mpc}^{-1}</math>)</i> . . . . .	20
4.1	<i>Contours représentant dans le cas flat les distributions posterior pour la comparaison entre notre étude et le survey EUCLID</i> . . . . .	31
4.2	<i>Contours représentant dans le cas non-flat les distributions posterior pour la comparaison entre notre étude et le survey EUCLID</i> . . . . .	32

# Liste des Tables

3.1	<i>Spécifications générales pour le survey spectroscopique DESI.</i> . . . . .	13
3.2	<i>Contraintes dans le cas population unique "fusion de colonnes" - flat - avec un modèle cosmologique (<math>w_0, w_a</math>).</i> . . . . .	18
3.3	<i>Contraintes dans le cas population unique "fusion de colonnes" - non-flat - avec un modèle cosmologique (<math>w_0, w_a</math>).</i> . . . . .	18
3.4	<i>Contraintes dans le cas multi-population - non-flat - avec un modèle cosmologique (<math>w_0, w_a</math>).</i> . . . . .	19
3.5	<i>Paramètres globaux pris sur l'ensemble des 3 populations BGS, LRG, ELG. Pour chaque colonne est indiqué le biais et la densité selon <math>z</math> (en <math>\text{deg}^{-2}</math>), la dernière correspondant au growth rate <math>f(z)</math>.</i> . . . . .	21
3.6	<i>Paramètres pris pour les émetteurs <math>H\alpha</math> du survey DESI dans le cas population unique. La première colonne indique le type d'objet, la deuxième le bin central du redshift, Les deux suivantes indiquent le biais et la densité selon <math>z</math>, et la dernière le growth rate <math>f(z)</math>.</i> . . . . .	21
3.7	<i>Paramètres de cross-corrélation sur les bins communs entre les populations LRG et ELG</i> . . . . .	22
4.1	<i>Spécifications pour le survey LSST.</i> . . . . .	24
4.2	<i>FoM (Figure of Merit) pour les 8 cas possibles de configuration.</i> . . . . .	28
4.3	<i>Synthèse des FoM pour (<math>GC_{\text{phot}} + \text{WL}</math>) versus (<math>GC_{\text{phot}} + \text{WL} + \text{XC}</math>) dans les cas (Y1, Y10) et (Flat, Non-Flat).</i> . . . . .	29
4.4	<i>Synthèse des surveys DESI et LSST avec toutes les sondes et différentes combinaisons (Flat/Non-Flat, (<math>S = \text{population unique} - \text{fusion de colonnes}</math>, <math>W = \text{population unique} - \text{moyenne pondérée}</math>, <math>M = \text{Multi populations}</math>) et Y1/Y10)</i> . . . . .	30



---

# Introduction

Découverte par Edwin Hubble en 1929, l'expansion de l'Univers fait toujours l'objet d'actives recherches. Par expansion, nous signifions que les galaxies s'éloignent d'autant plus rapidement qu'elles sont distantes de nous. Depuis une vingtaine d'années, les astrophysiciens ont constaté une autre particularité qui est l'accélération de cette expansion.

L'origine de l'expansion seule demeure encore aujourd'hui assez mal comprise même si le modèle de référence, appelé aussi "modèle standard" ou modèle  $\Lambda$ CDM, donne un cadre rendant compte de ce phénomène. Parmi ceux-ci est introduite une composante qui serait responsable de cette expansion. Dans une première approche, nous pourrions penser que l'Univers, régi par l'attraction gravitationnelle, s'effondrerait peu à peu sur lui-même. Or, il se passe la situation inverse et c'est précisément cette composante dénommée "Énergie noire" ou "Dark Energy" qui pourrait contrecarrer la tendance naturelle à ce que tous les corps s'attirent les uns avec les autres d'une manière globale.

Cette pression agit gravitationnellement de façon répulsive si elle est négative, produisant ainsi une force s'opposant à la gravitation de Newton. Dans les modèles les plus simples, l'équation d'état de cette composante s'exprime par une relation linéaire entre sa densité d'énergie et sa pression, avec un coefficient négatif.

Nous entrons désormais dans l'ère d'une cosmologie de précision où les incertitudes sur les paramètres cosmologiques ont une grande importance sur le modèle à privilégier et les contraintes sur l'équation d'état de l'énergie noire font partie de ces valeurs critiques.

Le but de cette étude est d'estimer, à partir du modèle standard, les contraintes que l'on peut espérer atteindre, avec un niveau de confiance donné, pour les prochaines collectes de données des deux télescopes au sol DESI et LSST. Pour cela, nous utiliserons le formalisme de Fisher qui est en lien avec une théorie plus générale qui est celle de la théorie de l'information. Cette méthode théorique, pouvant être implémentée numériquement, permet de mettre en relation l'ensemble des paramètres cosmologiques qui nous intéressent et de déterminer la distribution de chacun de ces paramètres en fonction des autres, ce qui peut faire apparaître ou non des corrélations.

Après avoir présenté de manière générale ce formalisme, nous verrons l'aspect pratique d'utilisation dans la partie spectroscopique (Galaxy clustering) pour le télescope DESI et pour LSST, les sondes photométriques (galaxy clustering (GC phot) et le "Lentillage faible" (Weak Lensing)). Nous en déduirons les différentes performances attendues et effectuerons une comparaison en combinant l'ensemble de ces contraintes dans le but de faire une synthèse des résultats.

# Modèle standard $\Lambda$ CDM

## 1.1 Équations de Friedmann-Lemaître

En assimilant la variable  $R(t)$  au rayon de l'Univers (que l'on appelle aussi "facteur d'échelle" suivant les conventions adoptées), celui-ci est gouverné par les équations, dites de Friedmann, suivantes :

$$\frac{R''}{R} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3} \quad (1.1)$$

avec  $G$  la constante de gravitation,  $\rho$  la densité de toutes les composantes présentes,  $p$  la pression du fluide cosmique et  $\Lambda$  la constante cosmologique.

La seconde équation s'écrit :

$$\left(\frac{R'}{R}\right)^2 = \frac{8\pi G\rho}{3} + \frac{\Lambda}{3} - \frac{kc^2}{R^2} \quad (1.2)$$

Ici,  $R(t)$  a la dimension d'une longueur suite au facteur  $c^2$  dans le terme  $-\frac{kc^2}{R^2}$  de (1.2). Le paramètre  $k$  est lié au type de géométrie considéré ( $k = 0$ , **univers plat (flat)**,  $k = \pm 1$ , **univers non plat (non-flat)**).

On rajoute quelques fois l'équation issue de la thermodynamique,  $\frac{dU}{dt} = -p\frac{dV}{dt}$  qui s'écrit simplement :

$$\frac{d(\rho c^2 R^3)}{dt} = -p\frac{dR^3}{dt} \quad (1.3)$$

## 1.2 Énergie noire dans le modèle classique

En reliant la pression de l'énergie noire à la densité par la relation :

$$p_{de} = w \rho c^2 \quad (1.4)$$

avec  $w$  un paramètre constant, on peut déduire, grâce à l'équation (1.3), l'évolution des différentes composantes du contenu de l'Univers. Un calcul simple nous donne pour la composante de densité  $\rho_i$  :

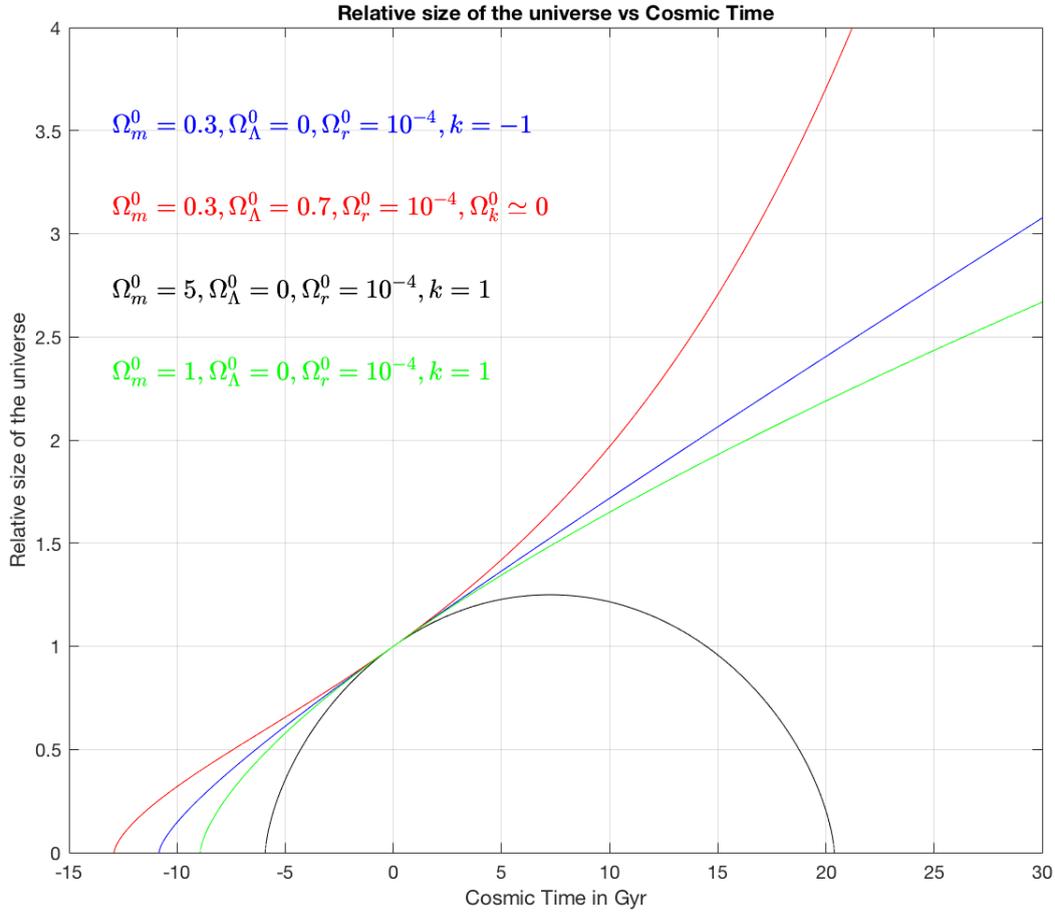
$$\rho_i = \rho_{i,0} \left[\frac{R_0}{R}\right]^{3(1+w)} = \rho_{i,0} (1+z)^{3(1+w)} \quad (1.5)$$

Avec ces relations, on retrouve la dépendance en  $R^{-3}$  pour la matière ordinaire ( $w = 0$ ),  $R^{-4}$  pour le rayonnement ( $w = 1/3$ ) et la pression  $p_{de} < 0$  pour l'énergie noire (si  $w < 0$ ), avec une valeur constante pour  $w = -1$ .

## 1.3 Histoire de l'expansion de l'Univers

En tenant compte de toutes les composantes contribuant à l'énergie (matière baryonique, matière noire, rayonnement, constante cosmologique), nous pouvons résoudre numériquement les équations

de Friedmann-Lemaître (1.1) et (1.2). Nous retraçons alors l'histoire cosmique de la dynamique de l'univers et notamment l'évolution de son facteur d'échelle. Voici ci-dessous les différents scénarios envisagés :



**FIGURE 1.1 – Facteur d'échelle normalisé de l'Univers en fonction du temps cosmique (1 Gyr = 1 Milliard d'années)**

Le modèle actuel qui correspond le mieux aux dernières observations est celui représenté par la courbe rouge : c'est le modèle  $\Lambda$ CDM ( $\Lambda$  représentant la constante cosmologique d'une valeur environ égale à 70% du contenu en énergie de l'Univers et l'acronyme CDM signifiant "Cold Dark Matter" qui représente une fraction de 25% : à noter que la matière classique dite baryonique représente approximativement les 5% restants. Les propriétés de la matière CDM, qualifiée de "froide" car non-relativiste, sont encore mal connues mais elle jouerait un rôle important dans la formation des premières grandes structures en servant de fixateur pour la matière baryonique le long des filaments cosmiques à grande échelle. Elle explique aussi le plateau que l'on observe sur les profils de vitesse dans les rotations des galaxies mais elle est rajoutée a posteriori, sans vraiment comprendre sa nature profonde.

Nous pouvons voir sur la figure ci-dessus que la courbe rouge présente une légère convexité au temps cosmique  $t = 0$ , c'est-à-dire à l'instant présent, ce qui traduit l'accélération actuelle de l'expansion, mise en évidence pour la première fois en 1998 par l'étude des courbes de lumière des supernovae SNIa (Perlmutter et al., 1998) et confirmée depuis par d'autres observations.

## 1.4 Extension de l'équation d'état de l'Énergie noire

Nous avons vu dans l'équation (1.4) que l'énergie noire admet comme un paramètre  $w = -1$  dans son équation d'état de base. Des modèles théoriques récents comme ceux de (Chevallier et Polarski, 2000), (Linder, 2003) et (Robert Scherrer, 2015) suggèrent une dépendance en redshift de l'équation d'état de l'énergie noire. Ceci se fait en introduisant 2 paramètres,  $w_0$  et  $w_a$ , apparaissant de cette manière :

$$w(z) = w_0 + w_a \frac{z}{1+z} \quad (1.6)$$

Le paramètre  $w_0$  correspond au paramètre  $w = -1$  décrit dans la section précédente, c'est-à-dire à la partie constante, tandis que  $w_a$  représente le facteur de la partie dépendante en redshift : cette paramétrisation correspondrait alors à une extension du modèle de concordance  $\Lambda$ CDM.

Ces 2 paramètres seront estimés dans notre étude et font l'objet d'un des résultats les plus attendus car leur valeur infirmerait ou non, avec un intervalle de confiance fixé, la validité de la relation (1.6). Cependant, nous verrons par la suite que d'autres paramètres cosmologiques ont aussi un rôle important dans l'évaluation des performances espérées pour les surveys DESI et LSST.

---

# Forecasting

## 2.1 Principe

Comme dit dans l'introduction, l'objectif du forecast est d'évaluer les performances d'un télescope, spatial ou au sol, avant sa mise en fonction opérationnelle. C'est pour cela que cette étude fait partie de la phase de design car elle intègre les caractéristiques futures du survey effectué. Il ne faut pas confondre les performances évoquées dans notre travail avec les performances techniques directement rattachées aux capacités instrumentales du télescope : les 2 sont évidemment liées mais dans le cadre de notre étude, nous nous attachons à l'évaluation de paramètres cosmologiques.

Pour obtenir ces propriétés propres au survey considéré, nous devons prendre en compte des paramètres déterminés avant la mission, que l'on qualifie de spécifications, comme par exemple la couverture totale du ciel observé, les différents types de populations des galaxies observées, leur densité respective en fonction du redshift, leur biais (c'est-à-dire la relation entre leur contraste de densité et le contraste de densité de la matière noire), ou encore les paramètres du modèle de référence choisi qui va notamment permettre de générer une distribution de matière fictive mais réaliste par rapport à nos contraintes de départ. Nous parlerons plus en détail de cet aspect dans la section décrivant la méthode appliquée au "Galaxy Clustering", technique qui modélise la distribution 3D des galaxies.

## 2.2 Formalisme de Fisher

Afin d'estimer les paramètres cosmologiques qui nous concernent, nous utilisons le formalisme de Fisher. L'application de cette théorie générale, adaptée à de nombreux problèmes en physique, a été utilisée pour la première fois en Cosmologie dans le papier de (Tegmark, 1997).

Cette technique issue de la statistique permet de tirer de l'information sur un paramètre dont on cherche à estimer la distribution "posterior", c'est-à-dire la distribution du paramètre concerné étant donné les valeurs observées prises en compte : elle est souvent notée  $p(\theta|d)$  avec  $\theta$  le paramètre et  $d$  les données. Pour faire une analogie avec d'autres grandeurs que l'on retrouve en statistiques, on assimile souvent la notion de matrice de Fisher au concept d'entropie de Shannon car elle permet de quantifier l'incertitude que l'on a sur des observables.

Cette distribution et les distributions associées avec les autres paramètres à évaluer sont déterminés en utilisant une fonction de score que l'on appelle aussi une fonction de vraisemblance et couramment notée  $\mathcal{L}$  ("Likelihood function") : comme nous le voyons ci-après dans la relation (2.1), la fonction likelihood est proportionnelle à la distribution "posterior"; nous pouvons donc écrire :  $\mathcal{L} = \mathcal{L}(d|\theta)$ . Dans le cas d'un paramètre unique, son utilisation permet de trouver la meilleure valeur de l'incertitude par rapport aux données.

Dans le cas multi-dimensionnel, le principe du calcul est le même mais nous nous servons d'une matrice appelée matrice de Fisher qui représente à la fois l'information de chaque variable aléatoire mais surtout leur inter-dépendance qui peut mettre en évidence des possibles relations mutuelles (on dit aussi corrélations).

Pour rappel, en inférence Bayésienne, la posterior  $p(\theta|d)$  se combine avec la likelihood  $\mathcal{L}(d|\theta)$  sous la forme générale :

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}} \quad (2.1)$$

où le facteur "prior" représente la probabilité d'obtenir le paramètre (en pratique, on prend une distribution uniforme) et l'évidence est utilisée comme facteur de normalisation vis-à-vis de la posterior. D'une manière plus formelle, cette relation ci-dessus se réécrit :

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)} \quad (2.2)$$

Dans notre étude, nous allons calculer les distributions posterior et croisées grâce à la méthode du maximum de vraisemblance, qui est un problème classique de maximisation. Concrètement, en considérant la fonction de vraisemblance  $\mathcal{L}$ , l'élément de la matrice de Fisher  $F_{\alpha\beta}$  est défini comme la valeur moyenne des dérivées secondes de son logarithme  $\ln\mathcal{L}$ , soit :

$$F_{\alpha\beta} = \left\langle \frac{-\partial^2 \ln\mathcal{L}}{\partial\theta_\alpha \partial\theta_\beta} \right\rangle, \quad (2.3)$$

où les indices  $\alpha$  et  $\beta$  représentent les indices des paramètres cosmologiques  $\theta_\alpha$  et  $\theta_\beta$  dont on se préoccupe. En supposant une fonction likelihood gaussienne, et une fois la matrice de Fisher calculée, la matrice de covariance de l'ensemble des variables est égale à l'inverse de la matrice de Fisher :

$$C_{\alpha\beta} = (\mathbf{F}^{-1})_{\alpha\beta} \quad (2.4)$$

Nous voyons alors l'intérêt majeur du calcul de la matrice de Fisher : en effet, une simple inversion matricielle suffit à obtenir des contraintes sur les grandeurs (la diagonale contient les variances et les termes non-diagonaux les covariances).

## 2.3 Critère de performances : Figure of Merit

Le but des prochains surveys DESI et LSST, tout comme le télescope EUCLID, est d'apporter une meilleure compréhension de l'Énergie noire et plus particulièrement de s'intéresser à son équation d'état, révélant ainsi ses propriétés physiques qui permettront d'étendre le modèle de concordance actuel. Dans l'équation (1.6) ont été introduits 2 paramètres cruciaux :  $w_0$  et  $w_a$ .

Afin d'évaluer les performances des 2 surveys, nous introduisons un critère, appelé "Figure of Merit" (FoM) qui quantifie la précision avec laquelle les paramètres de l'énergie noire sont contraints (voir (Albrecht et al., 2006)). Dans un cas standard, avec des likelihood supposées gaussiennes, cette FoM est directement issue de la matrice de Fisher et elle se réfère à l'inverse de l'aire délimitée par les contours à  $2\sigma$  (95% de niveau de confiance) des 2 paramètres considérés.

Dans notre travail, nous voulons mesurer la précision des contraintes obtenues sur les 2 variables  $w_0$  et  $w_a$ . Nous pouvons faire le lien de cette aire à  $2\sigma$  en prenant seulement en compte le bloc  $(\theta_\alpha, \theta_\beta) = (w_0, w_a)$  de la matrice de covariance. La FoM est alors égale à l'inverse de la racine carrée du déterminant de cette sous-matrice de taille 2x2. Nous avons ainsi la définition ci-dessous, avec la sous-matrice de covariance  $\widetilde{\text{Cov}}_{w_0, w_a}$  :

$$\text{FoM}_{w_0, w_a} = \frac{1}{\sqrt{\det(\widetilde{\text{Cov}}_{w_0, w_a})}} \quad (2.5)$$

## 2.4 Code de Boltzmann

Un forecast ne dispose pas de données expérimentales issues d'observations : nous allons donc les générer à partir d'un modèle de référence (appelé en anglais "fiducial") pour les paramètres cosmologiques. Un code de ce type s'appelle un code de Boltzmann : il va générer des spectres de puissance matière (qualifiés couramment de "P(k) matière") qui seront définis dans la section 3.6.

Les paramètres initiaux que nous prenons sont ceux du modèle de concordance ( $\Lambda$ CDM). Nous utiliserons comme code de Boltzmann le code CAMB (Lewis PA., 2000) (Code for Anisotropies in Microwave Background) disponible sur <https://camb.info/>. Ses nombreuses options et spécifications conviennent très bien à nos codes de prévisions. Dans toute notre étude, nous nous sommes servis de la version CAMB-0.1.7.

# Forecasting sur le survey DESI

## 3.1 Spécifications

Ma première tâche a été de rechercher et d'étudier les spécifications du survey DESI, afin de les intégrer dans les codes de Forecast existants. Le survey DESI (Dark Energy Spectroscopic Instrument) va collecter, dans la bande optique, environ 30 millions de spectres de galaxies et de quasars (quasi-stellar object) présélectionnés. Le but est de reconstruire l'histoire de la distribution de la matière l'Univers à travers une carte 3D jusqu'à un temps de regard en arrière de 11 milliards d'années selon le type d'objet observé. Ce survey récoltera aussi des données Lyman- $\alpha$  mais nous ne les utiliserons pas dans le cadre de notre étude.

Voici les principales caractéristiques du survey :

	Paramètre	Valeur
Surface du Survey dans le ciel	$A_{\text{survey}}$	14,000 deg <sup>2</sup>
Erreur de redshift spectroscopique	$\sigma_z$	0.0005(1 + z)
Redshift minimal et maximal des échantillons	$[z_{\text{min}}, z_{\text{max}}]$	[0.05 (BGS), > 2.1 (QSO)]

TABLE 3.1 – *Spécifications générales pour le survey spectroscopique DESI.*

Ce survey analysera 4 types d'objets rangés ci-dessous dans le sens croissant de redshift, c'est-à-dire selon une distance croissante :

- 1 : Les galaxies lumineuses dites "**BGS**" (**B**right **G**alaxies) : elles seront observées pour  $z < 0.4$  avec une magnitude apparente qui peut aller jusqu'à 20. Ces cibles sont qualifiées de faciles car elles sont nombreuses et permettent de s'intéresser à l'Univers récent, là où l'expansion est la plus forte.
- 2 : Les galaxies lumineuses rouges "**LRG**" (**L**uminous **R**ed **G**alaxies) observées pour  $z < 1$ . Ce sont les galaxies les plus massives et composées majoritairement de vieilles étoiles.
- 3 : Les galaxies avec émission de raies "**ELG**" (**E**mission **L**ine **G**alaxies) : elles représentent l'échantillon le plus important du survey. Elles sont plus éloignées que les **BGS** et **LRG** mais leur formation stellaire active permet de les détecter jusqu'à un redshift de  $z < 1.6$
- 4 : Les **QSO** (**Q**uasi **S**tellar **O**bjets) ou **Quasars** : ils correspondent à des galaxies contenant en leur centre un trou noir supermassif qui accrete une large quantité de matière, émettant ainsi un fort rayonnement. DESI sera capable de les détecter jusqu'à un redshift de  $z = 2.1$  ou plus.

## 3.2 Galaxy clustering spectroscopique

Nous décrivons ici brièvement la méthode appliquée pour la sonde spectroscopique "Galaxy clustering", en rappelant quels éléments constituent la matrice de Fisher et l'expression de la quantité primordiale pour notre forecast qui est le "**spectre de puissance observable**", noté  $P_{\text{obs}}$ .

### 3.3 Paramètres cosmologiques finaux à contraindre

Dans le cadre de notre travail et du formalisme de Fisher, nous devons choisir un ensemble de paramètres à estimer. Notre matrice de Fisher est construite initialement avec un jeu de paramètres décrit dans la section 3.5. Une fois la matrice construite avec ces paramètres initiaux, il nous faut choisir les paramètres finaux dont on veut obtenir les contraintes. Ici, nous prenons le même ensemble que celui du survey EUCLID (Euclid Collaboration et al., 2018), c'est-à-dire, dans l'ordre suivant :

$$\theta_{\text{final}} = \{\Omega_{\text{b},0}, h, \Omega_{\text{m},0}, n_{\text{s}}, \Omega_{\text{DE},0}, w_0, w_a, \sigma_8\} \quad (3.1)$$

$\Omega_{\text{b},0}$  représente le paramètre de densité de la matière baryonique ( $= \rho_0/\rho_c$ ),  $h$  la constante de Hubble réduite  $= H_0/100 \text{ km/s/Mpc}$ ,  $\Omega_{\text{m},0}$  la somme des paramètres de densité de la matière baryonique et CDM (Cold Dark Matter),  $n_{\text{s}}$  l'indice spectral qui caractérise l'invariance d'échelle du spectre de puissance,  $\Omega_{\text{DE},0}$  le paramètre réduit actuel de la constante cosmologique,  $w_0$  et  $w_a$  les paramètres dont nous avons parlé dans 1.4 avec l'équation (1.6),  $\sigma_8$  l'écart-type de l'amplitude des fluctuations de matière dans une sphère de  $R = 8 h^{-1} \text{ Mpc}$  à  $z = 0$ .

### 3.4 Projection

Notre matrice de Fisher va permettre de contraindre les paramètres cosmologiques  $\theta_{\text{final}}$  listés ci-dessus en (3.1). Pour obtenir ces contraintes à partir des paramètres de départ choisis dans la section 3.5 (notés  $\theta_{\text{initial}}$ ), nous devons faire un changement de coordonnées d'un point de vue matriciel et cette opération se fait grâce au calcul du Jacobien entre les 2 ensembles de grandeurs (B. Wang et al., 2007) et (Y. Wang, Percival et al., 2010). La matrice de Fisher  $S$  des paramètres  $\theta_{\text{final}}$  est reliée à la matrice originale  $F$ , c'est-à-dire construite avec les paramètres  $\theta_{\text{initial}}$  définis en (3.3), par l'expression analytique :

$$S_{ij} = \frac{\partial \theta_{\text{initial},\alpha}}{\partial \theta_{\text{final},i}} F_{\alpha\beta} \frac{\partial \theta_{\text{initial},\beta}}{\partial \theta_{\text{final},j}} \quad (3.2)$$

où les éléments  $\frac{\partial \theta_{\text{initial},\alpha}}{\partial \theta_{\text{final},i}}$  du Jacobien permettent cette nouvelle paramétrisation.

L'avantage de cette opération de projection est qu'une fois la matrice de Fisher construite avec le jeu  $\theta_{\text{initial}}$ , nous pouvons estimer tous les paramètres qui nous intéressent, à condition de pouvoir calculer les Jacobiens correspondants qui font le lien entre les paramètres de départ et d'arrivée.

### 3.5 Jeu de paramètres choisi pour la construction de la matrice de Fisher

Pour effectuer l'analyse spectroscopique "Galaxy clustering", autrement dit l'étude de la distribution des galaxies en 3D, nous suivons le même procédé que celui de l'IST de la mission EUCLID. La matrice de Fisher est produite avec l'ensemble des paramètres suivants, noté  $\theta_{\text{initial}}$ , en tenant compte de l'ordre :

$$\begin{aligned} \text{"Shape parameters"} & : : \{\omega_{\text{b}}, h, \omega_{\text{m}}, n_{\text{s}}\}^1, \\ \text{"Paramètres de nuisance"} \text{ non linéaires} & : : \{\sigma_{\text{p}}, \sigma_{\text{v}}\}, \\ \text{Paramètres "redshift dépendants"} & : : \{\ln D_{\text{A}}(z_i), \ln H(z_i), \ln f\sigma_8(z_i), \ln b\sigma_8(z_i), P_{\text{s}}(z_i)\}. \end{aligned} \quad (3.3)$$

Le paramètre  $\omega_{\text{b}}$  est défini par  $\omega_{\text{b}} = \Omega_{\text{b},0} h^2$  et  $\omega_{\text{m}}$  par  $\omega_{\text{m}} = \Omega_{\text{m},0} h^2$ . Les quantités  $\{\sigma_{\text{p}}, \sigma_{\text{v}}\}$  désignent les vitesses péculières : elles indiquent la vitesse locale d'une galaxie relativement à la vitesse globale issue de la loi d'expansion de Hubble-Lemaître. Elles sont fixées à des valeurs constantes dans notre analyse. Nous reviendrons sur ces 2 variables dans l'Appendice A car elles interviennent dans les anisotropies du spectre de puissance.

1. Nous omettons l'indice 0 dans les paramètres  $\omega$  pour avoir une notation plus légère

Les "Shape parameters" sont qualifiés ainsi car ils influent directement sur l'allure du spectre de puissance  $P_k$ . Les paramètres de nuisance non linéaires  $\sigma_p$  et  $\sigma_v$  correspondent à l'incertitude sur les effets de distorsions spatiales en redshift (**RSD**) qui modifient le signal des oscillations acoustiques baryoniques (**BAO**) : ceci est plus communément appelé l'effet "**Fingers-of-God**". De plus, les paramètres "redshift dépendants" sont pris en compte : il apparaît la distance angulaire, la constante de Hubble, le paramètre caractérisant l'effet **RSD** avec  $f\sigma_8(z) = f(z)\sigma_8(z)$ . Les 2 derniers paramètres sont le biais  $b\sigma_8(z) = b(z)\sigma_8(z)$  et le "shot noise residual"  $P_s(z)$  (bruit résiduel) : ils traduisent respectivement le biais sur les galaxies cibles et le bruit de fond résiduel. Ce "shot noise"  $P_s$  est dû au fait que la densité de galaxies est finie et donc ce bruit change avec la densité : il ne doit pas être confondu avec la notion de variance cosmique qui est liée au nombre de réalisations dans un processus aléatoire ; étant donné que nous n'avons qu'un seul Univers, nous n'avons qu'une seule réalisation : ceci impose une limite sur la capacité des surveys sur les contraintes que l'on peut espérer obtenir. Pour modéliser ce bruit "shot noise"  $P_s$ , nous prenons une distribution classique de Poisson indépendante de l'échelle, ce qui n'est pas tout à fait exact parce que les objets interagissent entre eux. Finalement, nous ajouterons  $P_s$  à l'expression du spectre de puissance observable  $P_{\text{obs}}$  (voir (3.7)).

Une de mes premières tâches a été de valider le modèle de biais  $b\sigma_8$  et de "shot noise"  $P_s$  qui sont utilisés dans les rapports techniques du survey DESI et les contraintes qui y sont rattachées.

### 3.6 Expression du spectre de puissance observable

L'évolution des fluctuations du contenu de l'Univers primordial peut-être décrit par des équations de fluide (Peebles, 1980). Pour une composante  $i$  donnée, ces fluctuations sont caractérisées par le contraste de densité suivant :

$$\delta_i(\vec{x}, z) \equiv \rho_i(\vec{x}, z) / \bar{\rho}_i(z) - 1, \quad (3.4)$$

qui mesure la déviation du champ de densité  $\rho_i(\vec{x}, z)$  par rapport au champ moyen  $\bar{\rho}_i(z)$  sur tout l'espace. Le vecteur  $\vec{x}$  représente les coordonnées spatiales à un redshift  $z$  donné. Pour étudier ces fluctuations d'un point de vue statistique, on décompose le contraste de densité sur la base de Fourier, ce qui se traduit par :

$$\delta_i(\vec{x}, z) = \int \frac{d^3k}{(2\pi)^3} \tilde{\delta}_i(\vec{k}, z) \exp(i\vec{k} \cdot \vec{x}). \quad (3.5)$$

Le spectre de puissance  $P_i(\vec{k}, z)$  est alors la transformée de Fourier inverse de la fonction de corrélation à 2 points et vérifie alors la relation :

$$\langle \tilde{\delta}_i(\vec{k}, z) \tilde{\delta}_i(\vec{k}', z) \rangle = (2\pi)^3 \delta_{\text{D}}(\vec{k} + \vec{k}') P_i(\vec{k}, z), \quad (3.6)$$

où  $\delta_{\text{D}}$  est la fonction delta de Dirac et  $\vec{k}$  est égal au vecteur nombre d'onde, c'est-à-dire à l'échelle considérée. Sous des conditions d'homogénéité et d'isotropie,  $|\vec{k}| = k$  et donc  $P_i(\vec{k}, z) = P_i(k, z)$ .

Notre prochaine étape est désormais de fournir l'expression analytique du spectre de puissance  $P_{\text{obs}}$ . Cette relation étant relativement complexe, nous devons résumer tous les effets dont elle tient compte :

- 1 : **le biais des galaxies** par rapport à la matière noire
- 2 : **les anisotropies dues aux distorsions spatiales** en redshift dont nous avons parlé dans la [section 3.5 \(RSD effect\)](#)
- 3 : **le bruit de fond** qui persiste malgré la suppression du bruit lié à la distribution de Poisson sur les galaxies cibles
- 4 : **l'incertitude sur les redshifts** qui introduit un bruit statistique
- 5 : **les effets de projection** : l'effet **Alcock-Paczynski (AP)** est responsable d'anisotropies dues à différentes renormalisations d'échelle dans les directions transversales et radiales.

Voici l'expression complète du spectre de puissance observable des émetteurs  $H_\alpha$ , incluant les anisotropies des spectres de puissance "matière" des galaxies :

$$P_{\text{obs}}(k, \mu; z) = \frac{1}{q_\perp^2 q_\parallel} \left\{ \frac{[b\sigma_8(z) + f\sigma_8(z)\mu^2]^2}{1 + [f(z)k\mu\sigma_p(z)]^2} \right\} \frac{P_{\text{dw}}(k, \mu; z)}{\sigma_8^2(z)} F_z(k, \mu; z) + P_s(z) \quad (3.7)$$

avec  $k$  égal au module du nombre d'onde et  $\mu$  le cosinus de l'angle entre la ligne de visée et la direction du vecteur d'onde  $\vec{k}$ . Les 4 nouvelles quantités apparaissant dans l'expression de  $P_{\text{obs}}$  sont  $q_\perp$ ,  $q_\parallel$ ,  $P_{\text{dw}}$  et  $F_z(k, \mu; z)$ . Pour résumer, ils représentent respectivement le ratio entre la distance angulaire réelle et de référence, le ratio entre le paramètre de Hubble réel et de référence,  $P_{\text{dw}}$  ("de-wiggled") correspondant au spectre de puissance qui tient compte de l'amortissement des BAO et finalement,  $F_z(k, \mu; z)$  nommé le "facteur exponentiel" qui inclut les erreurs possibles sur la mesure des redshifts le long de la ligne de visée. Le détail de l'obtention de cette relation peut-être trouvé dans l'Appendice A.

## 3.7 Cas population unique et multi-populations

### 3.7.1 Construction matrice de Fisher avec population unique

D'après les papiers (Tegmark, 1997) et (Seo et Eisenstein, 2003), on calcule l'élément  $F_{ij}$  de la matrice pour le spectre de puissance "matière" avec des bins de redshift centrés sur  $z_i$  comme ceci :

$$F_{\alpha\beta}^{\text{bin}}(z_i) = \frac{1}{8\pi^2} \int_{-1}^1 d\mu \int_{k_{\text{min}}}^{k_{\text{max}}} k^2 dk \left[ \frac{\partial \ln P_{\text{obs}}(k, \mu; z_i)}{\partial \alpha} \frac{\partial \ln P_{\text{obs}}(k, \mu; z_i)}{\partial \beta} \right] V_{\text{eff}}(z_i; k, \mu) \quad (3.8)$$

où  $\alpha$  et  $\beta$  sont les paramètres cosmologiques que l'on souhaite faire varier et dont nous avons donné la liste dans la section 3.5. Nous pouvons ensuite effectuer une projection pour calculer la matrice de Fisher des paramètres finaux énumérés dans la section 3.3.

Un point important concerne les paramètres indépendants du redshift (les "shape-parameters") où l'on doit sommer sur les différents bins  $z_i$  l'élément  $F_{ij}$ , ce qui n'est pas le cas pour les variables ayant une dépendance en  $z$ .

### 3.7.2 Construction matrice de Fisher avec multiples populations

Le survey DESI compte certains bins de redshift dans lesquels 2 populations se superposent : il faut donc construire une matrice de Fisher prenant en compte les corrélations croisées entre les différentes populations. D'après le papier (White, Song et Percival, 2008), l'élément  $F_{ij}$  de la matrice se calcule en faisant intervenir l'inverse de la **matrice de covariance des observables**  $C$  associée aux différents  $P_{\text{obs}}$ , c'est-à-dire selon :

$$F_{ij} = \sum_{XY} \int \frac{V_0 d^3k}{(2\pi)^3} \left( \frac{\partial P_{\text{obs},X}}{\partial p_i} \right) (C^{-1})_{XY} \left( \frac{\partial P_{\text{obs},Y}}{\partial p_j} \right) \quad (3.9)$$

où les indices  $(i, j)$  sont relatifs aux paramètres considérés et les indices  $X$  et  $Y$  représentent les indices des populations de galaxies prises en compte. Il faut préciser que dans notre cas, nous faisons la cross-corrélation entre 2 populations, ce qui fait que nous travaillerons avec une matrice de covariance  $C_{XY}$  de taille 3x3 (un bloc 2x2 pour les 2 auto-corrélations et une 3<sup>ème</sup> dimension pour les corrélations croisées). Le calcul de l'élément étant numérique, la valeur de l'intégrale nécessite une interpolation pour les spectres de Puissance "matière" issu du code CAMB. Étant donné que je dois faire le produit entre les composantes des 2 tableaux de dérivées et celui des matrices 3x3, j'ai codé, pour rendre les algorithmes plus simples à mettre en œuvre un tableau 4D pour  $C_{XY}$ , autrement dit une matrice contenant le nombre adéquat de matrices 3x3 : on peut alors inverser pour chaque composante de la grille 2D ces matrices de taille réduite.

La difficulté ici est de choisir, lors de la multiplication de "points à points" entre les 3 facteurs apparaissant dans l'intégrale, les bonnes dérivées du spectre de puissance observable par rapport au

jeu de paramètres initiaux et de sommer sur les bons indices concernant les éléments de la matrice inverse de covariance utilisée. Pour les éléments "Shape parameters" de la matrice ( $\omega_b, h, \omega_m, n_s$ ), il ne faut pas oublier de sommer sur tous les bins de redshift.

### 3.7.3 Cas Univers plat et non plat

Pour un Univers plat, on fixe le paramètre de courbure  $k = 0$  qui apparaît, pour un Univers FLRW (Friedmann-Lemaître-Robertson-Walker), sous la forme :

$$\Omega_m^0 + \Omega_r^0 + \Omega_{DE}^0 + \Omega_k^0 = 1 \quad (3.10)$$

Soit, avec  $\Omega_k^0 = 0$ , on déduit :  $\Omega_{DE}^0 = 1 - \Omega_m^0 - \Omega_r^0 = 1 - \Omega_{CDM}^0 - \Omega_b^0 - \Omega_r^0$ , où  $\Omega_{CDM}^0$  est la densité réduite actuelle de la matière noire froide (Cold Dark Matter).

Dans le cas non-flat, le paramètre de courbure est un paramètre libre, il va donc lui aussi dépendre de l'estimation des autres observables de densité. De plus, nous ne tenons pas compte dans notre code de  $\Omega_r^0$  car cela ne change presque pas les contraintes (moins de 0.1%). Enfin, les neutrinos ne sont pas non plus pris en compte : ils ont un impact minime sur les résultats d'estimations ( $< 2\%$ ).

### 3.7.4 Résultats unique-multi populations

#### 3.7.4.1 Méthode population unique - fusion de colonnes

Nous affichons ici les contours correspondant à un traitement "population unique avec fusion de colonnes". Son implémentation est décrite dans la section 3.8.1.1.

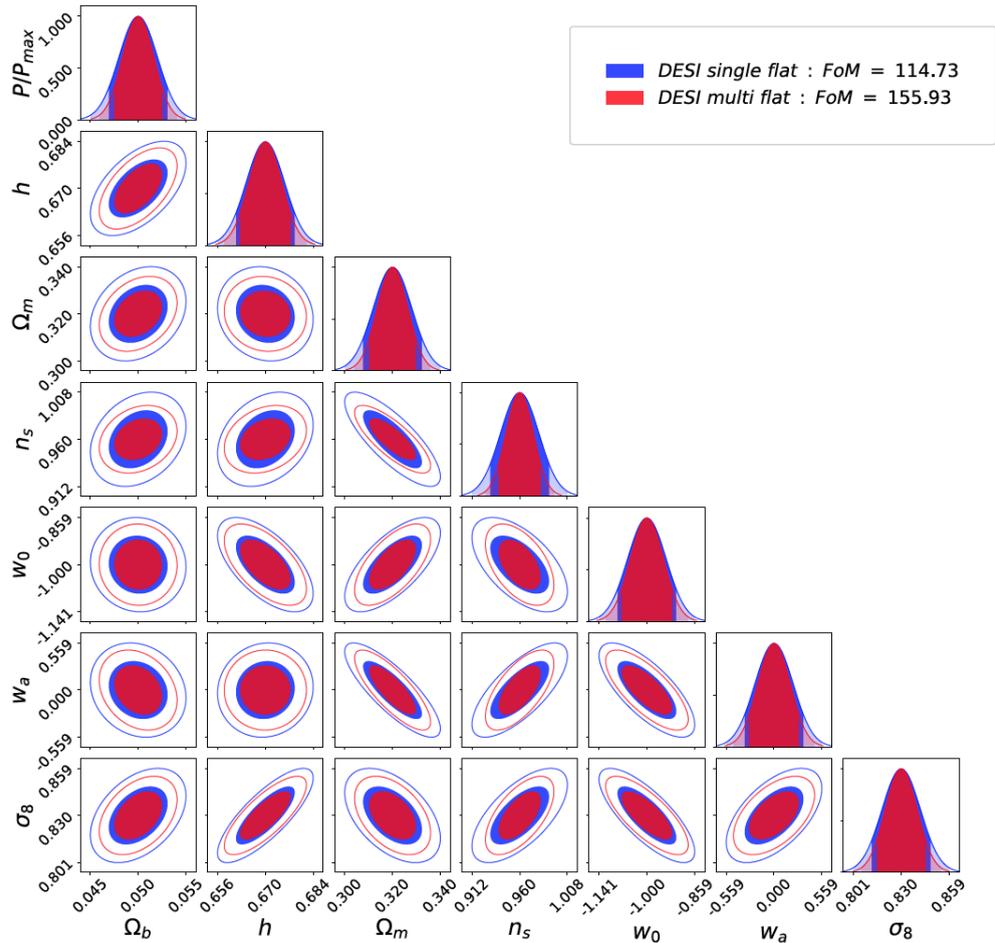


FIGURE 3.1 – Contours représentant les distributions posterior avec population unique "fusion de colonnes" et multi dans le cas flat ( $k_{\max} = 0.25 \text{ h.Mpc}^{-1}$ )

TABLE 3.2 – Contraintes dans le cas population unique "fusion de colonnes" - flat - avec un modèle cosmologique ( $w_0, w_a$ ).

Modèle ( $w_0, w_a$ ) CDM flat, single, merging FoM = 114.73							
	$\Omega_{b,0}$	$h$	$\Omega_{m,0}$	$n_s$	$w_0$	$w_a$	$\sigma_{8,0}$
Valeurs fiducial	0.05	0.67	0.32	0.96	-1.0	0.0	0.830024
Contraintes ( $\pm\sigma$ )	0.002210	0.005808	0.008067	0.019379	0.056881	0.225351	0.011848

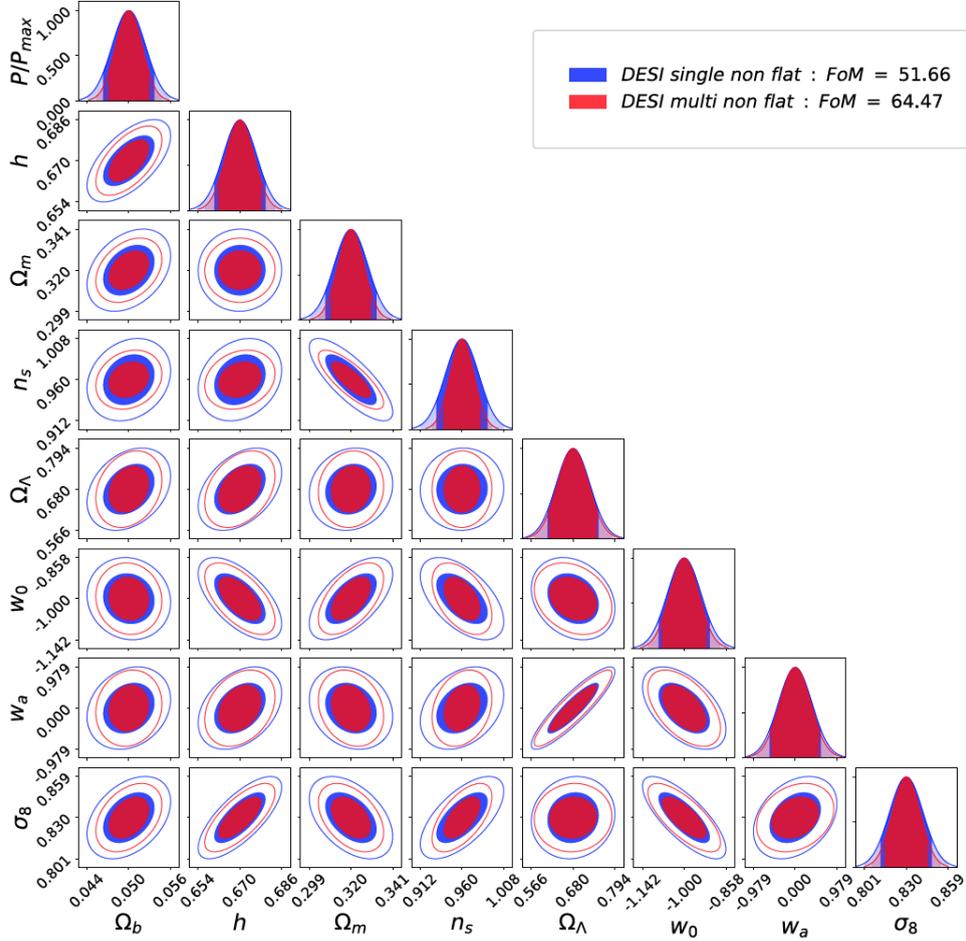


FIGURE 3.2 – Contours représentant les distributions posterior avec population unique "fusion de colonnes" et multi dans le cas non-flat ( $k_{\max} = 0.25 \text{ h.Mpc}^{-1}$ )

TABLE 3.3 – Contraintes dans le cas population unique "fusion de colonnes" - non-flat - avec un modèle cosmologique ( $w_0, w_a$ ).

Modèle ( $w_0, w_a$ ) CDM non-flat, single, merging FoM = 51.66								
	$\Omega_{b,0}$	$h$	$\Omega_{m,0}$	$n_s$	$\Omega_{DE,0}$	$w_0$	$w_a$	$\sigma_{8,0}$
Valeurs fiducial	0.05	0.67	0.32	0.96	0.68	-1.0	0.0	0.830024
Contraintes ( $\pm\sigma$ )	0.002375	0.006319	0.008444	0.019499	0.045970	0.057293	0.394452	0.011844

### 3.7.4.2 Méthode population unique - moyenne pondérée

Nous affichons ici les contours correspondant à un traitement "population unique avec moyenne pondérée". Son implémentation est décrite dans la section 3.8.1.2.

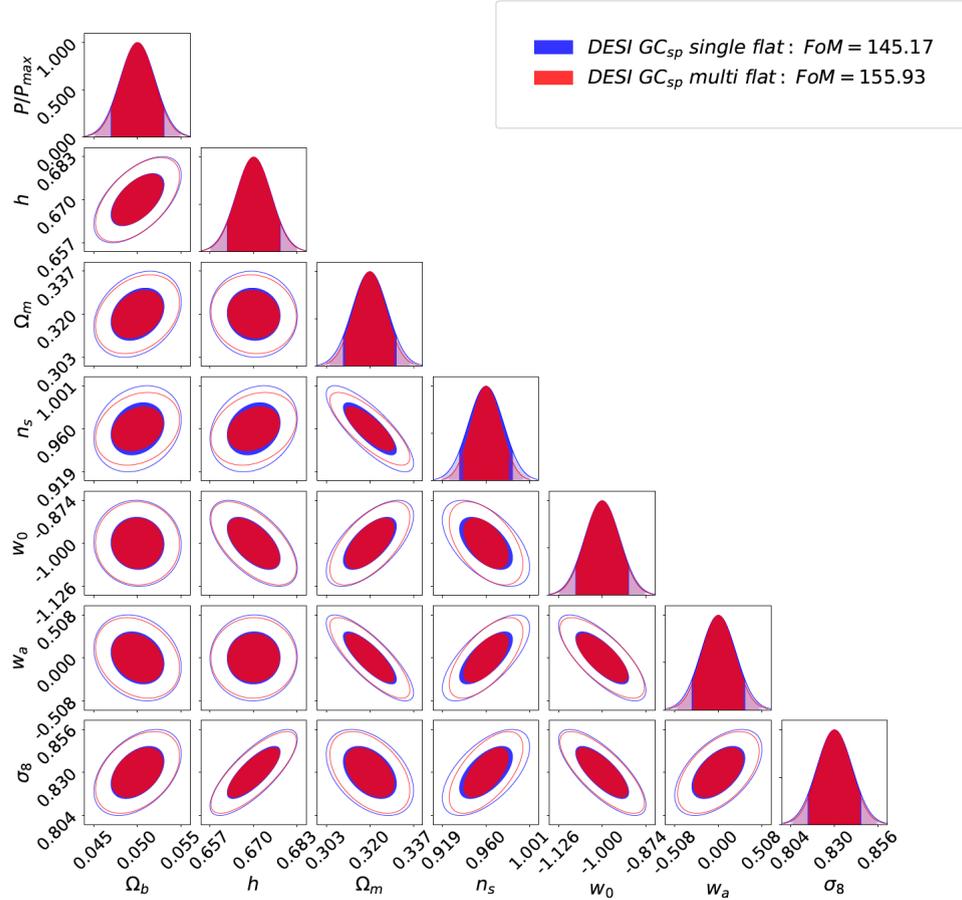


FIGURE 3.3 – Contours représentant les distributions posterior avec population unique "moyenne pondérée" et multi dans le cas flat ( $k_{\max} = 0.25 \text{ h.Mpc}^{-1}$ )

### 3.7.4.3 Interprétation des méthodes "single/multi populations"

TABLE 3.4 – Contraintes dans le cas multi-population - non-flat - avec un modèle cosmologique ( $w_0, w_a$ ).

Modèle ( $w_0, w_a$ ) CDM non-flat,multi FoM = 64.47							
$\Omega_{b,0}$	$h$	$\Omega_{m,0}$	$n_s$	$\Omega_{DE,0}$	$w_0$	$w_a$	$\sigma_{8,0}$
<b>Valeurs fiducial</b>							
0.05	0.67	0.32	0.96	0.68	-1.0	0.0	0.830024
<b>Contraintes (<math>\pm\sigma</math>)</b>							
0.001878	0.005325	0.006590	0.013994	0.042769	0.049821	0.365444	0.009847

On remarque immédiatement que les contraintes associées au cas flat sont meilleures que celles obtenues en non-flat, pour les 2 cas "single population" (fusion de colonne et moyenne pondérée) comme "multi" (par exemple,  $\text{FoM}_{\text{flat},\text{single},\text{merging}} = 114.73$  alors que  $\text{FoM}_{\text{non-flat},\text{single},\text{merging}} = 51.66$ ) : ceci est dû au fait que l'on fixe le paramètre de courbure  $\Omega_k$ , ce qui fait un paramètre de moins à évaluer et donc améliore les estimations des autres. On peut saisir intuitivement la différence entre ces 2 cas.

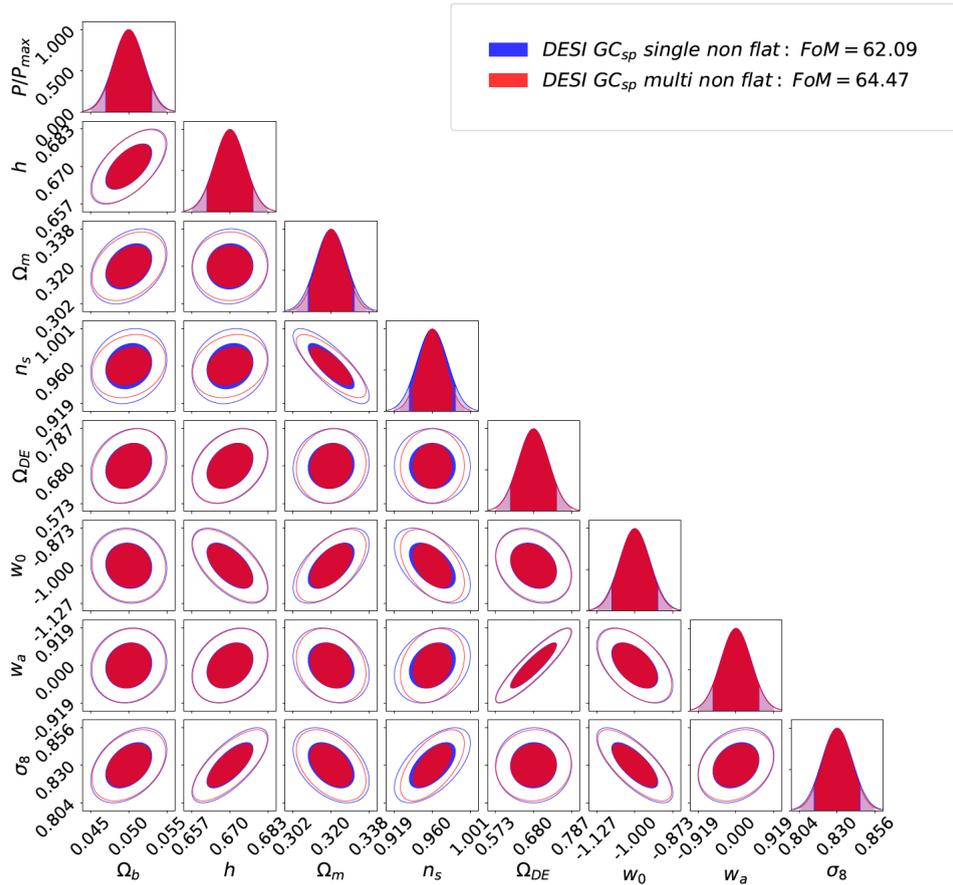


FIGURE 3.4 – Contours représentant les distributions posterior avec population unique "moyenne pondérée" et multi dans le cas non-flat ( $k_{\max} = 0.25 \text{ h.Mpc}^{-1}$ )

Une autre validation de ces résultats est acquise par le fait que les contraintes multi-populations sont meilleures que celles obtenues avec un seul type de galaxie : en effet, les cross-corrélations apportent de l'information supplémentaires sur les paramètres cosmologiques et en particulier sur le couple  $(w_0, w_a)$ , d'où une augmentation de la FoM pour le cas "multi" ( $\text{FoM}_{\text{non-flat, multi}} = 64.47$  par rapport au cas single "fusion de colonnes" où  $\text{FoM}_{\text{non-flat, single, merging}} = 51.66$ ). Nous voyons ici tout l'intérêt des multi-traceurs qui permettent de réduire notamment la variance cosmique. Le gain est d'environ 36% dans le cas flat et 25% dans le cas non flat.

À noter qu'avec un échantillon pauvre au niveau statistique (on dit aussi redondant ou d'entropie de Shannon faible), le gain sera moindre car la différence entre ajout d'échantillon et information non significative sera moins élevée.

## 3.8 Codage Python population unique et cross-corrélations

### 3.8.1 Implémentation population unique

Le module SpecSAF du code Totally SAF développé par **Safir Yahia-Cherif** (Doctorant IRAP, équipe GAHEC) permet de calculer des contraintes pour une population unique dans sa version de base.

Mon premier travail a été d'inclure le modèle de biais et les  $P_{\text{shot}}$  issus des spécifications de DESI. Pour cela, nous disposons dans le rapport technique DESI des données statistiques suivantes pour les trois types de galaxies considérées :

**TABLE 3.5** – *Paramètres globaux pris sur l'ensemble des 3 populations BGS, LRG, ELG. Pour chaque colonne est indiqué le biais et la densité selon  $z$  (en  $\text{deg}^{-2}$ ), la dernière correspondant au growth rate  $f(z)$ .*

Redshift $z_c$	BGS : biais( $z$ )	LRG : biais( $z$ )	ELG : biais( $z$ )	BGS : $dn/dz$	LRG : $dn/dz$	ELG : $dn/dz$	$f(z)$
0.175	1.1133849956	0.0000000000	0.0000000000	2862.3766865	0.0000000000	0.0000000000	0.63004698923
0.425	1.7983127401	0.0000000000	0.0000000000	204.12110815	0.0000000000	0.0000000000	0.73889083367
0.65	0.0000000000	1.4469899900	0.7149832900	0.0000000000	832.00000000	309.00000000	0.80866367407
0.85	0.0000000000	1.4194157200	0.7013583500	0.0000000000	662.00000000	1923.00000000	0.85346288717
1.05	0.0000000000	1.4006739400	0.6920977100	0.0000000000	51.0000000000	1441.00000000	0.88639317400
1.25	0.0000000000	0.0000000000	0.6856214000	0.0000000000	0.0000000000	1337.00000000	0.91074033422
1.45	0.0000000000	0.0000000000	0.6809754100	0.0000000000	0.0000000000	466.00000000	0.92892482272
1.65	0.0000000000	0.0000000000	0.6775659400	0.0000000000	0.0000000000	126.00000000	0.94267299703

**Remarque** : le growth rate peut être approximé par ( $f(z) = |\Omega_m(z)|^\gamma$ ) mais il est calculé dans notre code de manière beaucoup plus précise par la résolution de l'équation différentielle des perturbations.

### 3.8.1.1 Implémentation "single population" : fusion de colonnes

Pour faire en sorte d'avoir une population par bin de redshift avec un biais et un  $P_{\text{shot}}$  différents de 0, nous sélectionnons les 2 premiers bins de la population BGS et les 6 autres bins de la population ELG, ce qui constitue une population unique de 8 bins au total.

**TABLE 3.6** – *Paramètres pris pour les émetteurs  $H\alpha$  du survey DESI dans le cas population unique. La première colonne indique le type d'objet, la deuxième le bin central du redshift, Les deux suivantes indiquent le biais et la densité selon  $z$ , et la dernière le growth rate  $f(z)$ .*

Type objet	$z_c$	biais( $z$ )	$dn/dz$	$f(z)$
BGS	0.175	1.1133849956	2862.3766865	0.63004698923
BGS	0.425	1.7983127401	204.12110815	0.73889083367
ELG	0.65	0.7149832900	309.00000000	0.80866367407
ELG	0.85	0.7013583500	1923.00000000	0.85346288717
ELG	1.05	0.6920977100	1441.00000000	0.88639317400
ELG	1.25	0.6856214000	1337.00000000	0.91074033422
ELG	1.45	0.6809754100	466.00000000	0.92892482272
ELG	1.65	0.6775659400	126.00000000	0.94267299703

**Remarque** : nous ne tenons pas compte dans nos calculs des galaxies de type QSO car leur densité n'est pas assez élevée d'un point de vue statistique : en effet, elles sont surtout utilisées pour les raies d'absorption Lyman- $\alpha$  à haut redshift caractérisant le milieu intergalactique présent entre ces objets et nous.

### 3.8.1.2 Implémentation "single population" : moyenne pondérée

Le contraste de densité de galaxies  $\delta_g(\vec{r}, z)$  est relié à celui de la densité de matière noire  $\delta_{\text{DM}}(\vec{r}, z)$  par :

$$\delta_g(\vec{r}, z) = b(z) \delta_{\text{DM}}(\vec{r}, z) \quad (3.11)$$

où  $b(z)$  est appelé le biais de l'échantillon considéré.

La fonction de corrélation à 2 points  $\xi(\vec{r})$  étant la transformée de Fourier du spectre de puissance matière, on a la relation :

$$\xi_g(\vec{r}, z) = b(z)^2 \xi_{\text{DM}}(\vec{r}, z) \quad (3.12)$$

Maintenant, exprimons la somme des densités locales  $N_1$  et  $N_2$  de 2 échantillons de galaxies différents, en faisant intervenir les contrastes de densité définis en (3.11) et les densités globales moyennes  $n_1$  et  $n_2$  :

$$\delta_{g1} = b_1 \delta_{\text{DM}} = \frac{N_1 - n_1}{n_1} = \frac{N_1}{n_1} - 1 \Rightarrow N_1 = n_1 b_1 \delta_{\text{DM}} + n_1 \quad (3.13)$$

Formule identique pour le deuxième échantillon :

$$\delta_{g2} = b_2 \delta_{\text{DM}} = \frac{N_2 - n_2}{n_2} = \frac{N_2}{n_2} - 1 \Rightarrow N_2 = n_2 b_2 \delta_{\text{DM}} + n_2 \quad (3.14)$$

On en déduit d'après la somme  $N = N_1 + N_2$  :

$$N = N_1 + N_2 = \delta_{\text{DM}} (n_1 b_1 + n_2 b_2) + n_1 + n_2 = \delta_{\text{DM}} (n_1 b_1 + n_2 b_2) + n \quad (3.15)$$

En exprimant  $N$  avec le contraste de densité  $\delta_g$  et le paramètre de bias représentatif  $b$  :

$$N = n b \delta_{\text{DM}} + n \quad (3.16)$$

On peut conclure sur le biais caractérisant de manière statistique les 2 échantillons, pour un redshift  $z$  donné :

$$b(z) = \frac{1}{n} (n_1(z) b_1(z) + n_2(z) b_2(z)) = \frac{1}{n_1 + n_2} (n_1 b_1 + n_2 b_2) \quad (3.17)$$

Ce biais représente la moyenne pondérée des 2 biais par rapport à la densité globale de galaxies dans chacun des 2 échantillons. Dans notre étude, nous prenons comme première population la fusion des 2 premiers bins BGS, des 3 suivants de LRG et des 3 derniers de ELG. La seconde population est prise avec les 3 bins intermédiaires de ELG. On fait ainsi la moyenne pondérée sur les 3 bins qui s'enchevêtrent entre les populations LRG et ELG (pour  $z = 0.65, 0.85, 1.05$ ).

### 3.8.2 Implémentation cross-corrélations

La cross-corrélation consiste à regrouper des informations partagées par des mêmes bins de redshift. Dans notre cas, nous nous sommes aperçus qu'il y avait des propriétés que l'on pouvait regrouper pour les bins de redshift  $\mathbf{z}=(0.65,0.85,1.05)$ . Nous pouvons donc appliquer la formule du papier de White (3.9) précédemment évoquée.

Voici sur le tableau suivant un récapitulatif des bins corrélés :

**TABLE 3.7 – Paramètres de cross-corrélation sur les bins communs entre les populations LRG et ELG**

Type objet	$z_c$	biais(z)	dn/dz	$f(z)$
LRG	0.65	1.4469899900	832.00000000	0.80866367407
LRG	0.85	1.4194157200	652.00000000	0.85346288717
LRG	1.05	1.4006739400	51.00000000	0.88639317400
ELG	0.65	0.7149832900	309.00000000	0.80866367407
ELG	0.85	0.7013583500	1923.00000000	0.85346288717
ELG	1.05	0.6920977100	1441.00000000	0.88639317400

Etant donné que les éléments de Fisher sont cumulés sur les redshifts pour les paramètres de forme ("shape parameters"), le code que j'ai implémenté traite les redshifts non corrélés avec la formule classique (3.8) et ceux corrélés avec la formule de White (3.9). Pour les paramètres dépendants du redshift le calcul sera similaire mais sans cumuler les composantes de la matrice.

### 3.8.3 Optimisation GPU

La possibilité d'une optimisation sur la formule de White s'est relativement assez vite posée. Sa nécessité n'est pas liée à la taille des matrices à inverser (matrice 3x3) mais au grand nombre d'inversions à effectuer : on qualifie en informatique cette tâche de "**batch processing**". Cette situation se présente dans la partie cross-corrélée, où un nombre important de matrices 3x3 doivent être inversées et ceci de manière indépendantes entre elles. Ces tâches séparées conviennent parfaitement au calcul **GPU** qui est capable d'exécuter sur des milliers de threads des opérations similaires. Le listing du code de la partie **GPU PyCuda** (framework pour cartes graphiques **NVIDIA**) est disponible sur le lien suivant : <https://github.com/fabien-dournac/forecast-gpu-optimization/blob/master/gpu-code.py>.

Pour avoir un ordre de grandeur sur le **SpeedUp (runtime code séquentiel / runtime code parallélisé)**, cette optimisation permet de réaliser avec une carte NVIDIA standard, pour un même nombre d'appels de fonctions d'inversions (Numpy et PyCuda) sur une grille 2D 500x500 (soit 250.000 matrices 3x3), **l'inversion GPU en 1 seconde** environ contre **3 secondes pour la fonction `numpy.linalg.inv`**. Nous gagnons donc un facteur 3 sur le temps d'exécution de cette partie, ce qui fait, si l'on effectue une pondération avec 3 bins corrélés sur 8, un gain proche de 2 sur le runtime total. Cette valeur de SpeedUp n'est donc pas négligeable sur un aspect purement **HPC** (High Performance Computing) et permet donc d'obtenir des résultats plus rapidement. Il faut signaler qu'il existe l'équivalent multi-plateformes de **PyCuda** pour les cartes graphiques non NVIDIA comme par exemple les cartes **AMD** : c'est le framework **PyOpenCL** dont la syntaxe et la méthodologie sont très proches de celles de **PyCuda**.

### 3.8.4 Tâches principales et validation de l'implémentation

Mon travail a consisté en grande partie à adapter le code de forecast existant pour le relevé EU-CLID afin d'implémenter l'approche "multi-populations" pour le relevé DESI décrite dans la section précédente. Concernant les termes uniquement diagonaux de **la matrice de covariance des observables  $C$**  (3.9), une comparaison des résultats avec 2 approches différentes a permis de confirmer rapidement la conformité de mon programme. Pour le cas général, c'est-à-dire en incluant aussi les termes non-diagonaux, la tâche s'est révélée plus ardue mais la validation finale des résultats, avec une optimisation GPU, a pu être obtenue.

# Forecasting sur le survey LSST

## 4.1 Spécifications

Le Large Synoptic Survey Telescope (LSST) va faire partie d'une nouvelle génération de télescopes par ses dimensions et caractéristiques jamais atteintes pour un instrument au sol faisant du relevé astronomique. Voici ci-dessous ses principales spécifications :

	Paramètre	Valeur
Surface du Survey dans le ciel	$A_{\text{survey}}$	20,000 deg <sup>2</sup>
Erreur sur le redshift Weak Lensing (Y1 et Y10) :	$\sigma_z$	0.05(1 + z)
Erreur sur le redshift GC photométrique (Y1 et Y10) :	$\sigma_z$	0.03(1 + z)
Redshift minimal et maximal des échantillons (Y1 et Y10) :	$[z_{\text{min}}, z_{\text{max}}]$	[0.2, 1.2]
Nombre de bins : (Y1) : 5 ( $\Delta z = 0.2$ ); (Y10) : 10 ( $\Delta z = 0.1$ )		

TABLE 4.1 – *Spécifications pour le survey LSST.*

Ce télescope permettra aux astrophysiciens de faire du **Weak Lensing** (Lentillage Faible), autrement dit l'étude des déformations d'images de galaxies par les différentes zones de matière baryonique et de matière noire durant tout le long du trajet de la lumière jusqu'à nous. Il donnera aussi l'opportunité de faire du **Galaxy Clustering Photométrique** : cette technique consiste à déduire les propriétés cosmologiques de notre Univers (tout comme le Weak Lensing) à partir de la distribution photométrique de la matière dans le Cosmos. Cette méthode est semblable au Galaxy clustering classique mais avec la mesure des redshifts des galaxies en moins : c'est de l'imagerie dans la mesure où l'on analyse le flux bolométrique contrairement au GC où nous caractérisons des zones particulières du spectre. À priori, nous pourrions penser que nous aurons seulement une vision 2D (angulaire) avec le Weak Lensing mais en réalité, nous pouvons effectuer de la "tomographie" (c'est-à-dire reconstruire un volume à partir de plusieurs plans de coupe) avec des redshifts photométriques. L'abréviation Y1 signifie une année d'observation et Y10 10 ans.

## 4.2 Weak Lensing

Cette section décrit la technique de forecasting que nous utilisons pour le Weak Lensing (WL). Cette méthode repose une observable que l'on appelle le spectre de puissance tomographique de cisaillement (tomographic shear cosmic power spectrum). Le phénomène physique est le lentillage faible, c'est-à-dire la déformation des images des galaxies tandis que le "tomographic shear power spectrum" représente l'observable, autrement dit ce qui correspond à ce que l'on observe concrètement.

### 4.2.1 Caractéristiques du phénomène de Weak Lensing

Il y a 2 types de déformations que l'on considère habituellement dans le WL : la première provoque la **déformation de la taille de la galaxie et le changement de son intensité**, elle est qualifiée de **déformation convergente**. La seconde, appelée **distorsions de cisaillement (Shear distortions)** : elle induit un aplanissement et une modification de l'excentricité sous laquelle est vue la galaxie.

Nous considérons ici uniquement les distorsions de cisaillement : ce sont ses effets, par l'intermédiaire de son spectre de puissance issu de la fonction de corrélation à 2 points, dont nous pourrions déduire des informations sur les paramètres cosmologiques. Différentes manières de traiter le spectre de puissance tomographique existent mais nous retiendrons celle qui associe le spectre à une distribution angulaire. Ainsi, par rapport à l'analyse GC (Galaxy Clustering) du survey DESI, d'autres paramètres et phénomènes sont à considérer dans nos prévisions :

- 1 : **Le spectre de puissance cosmique de cisaillement théorique.**
- 2 : **Le spectre de puissance d'alignement intrinsèque** : il modélise l'alignement local des galaxies.
- 3 : **Les petites échelles du spectre de puissance matière** représentant les halos de matière noire ayant une influence majeure sur le spectre de puissance de cisaillement.
- 4 : **Les redshifts tomographiques et la densité des objets** qui modélise l'incertitude sur la position des galaxies due aux erreurs sur l'estimation de leurs redshifts.
- 5 : **Le "shot noise"** que nous avons déjà évoqué dans le survey DESI et qui est représenté par un bruit de Poisson.

Chacune de ces composantes joue un rôle dans la détermination de notre forecast. Il est difficile de résumer de manière précise tout le formalisme Fisher appliqué au Weak Lensing. C'est pourquoi nous nous attachons à énumérer les formules les plus importantes, permettant alors de mieux se concentrer sur la structure des codes existants et des spécifications propres au survey LSST qui produiront nos résultats. Pour plus de précisions, je renvoie le lecteur à [l'Appendice B](#).

## Spectre de puissance cosmique de cisaillement

Les effets de lentille gravitationnelle changent l'ellipticité des galaxies lors des observations par rapport à leur forme réelle. Ceci est dû à la présence de structures plus ou moins massives le long de la ligne de visée qui engendrent des arcs gravitationnels ou la multiplication d'images d'une seule galaxie (comme dans le célèbre exemple de la croix d'Einstein qui est obtenu par un lentillage fort).

Pour une galaxie seule, on définit cette modification d'ellipticité sous la forme :

$$\epsilon = \gamma + \epsilon^l \quad (4.1)$$

où  $\gamma$  est le cisaillement cosmologique et  $\epsilon^l$  l'ellipticité propre (non lentillée). L'ellipticité encode 2 types d'informations : sa valeur moyenne est considérée comme nulle sur une large surface du ciel alors qu'elle a une valeur non nulle pour la fonction de corrélation à 2 points. Elle contient à la fois des informations sur l'histoire de l'expansion de l'Univers ainsi que sur la distribution de ses structures à travers le spectre de puissance matière.

Le calcul du spectre de puissance angulaire dans sa version basique se fait classiquement par les harmoniques sphériques de la transformation de Legendre. Le calcul complet est laborieux (Taylor et al., 2018), mais une approximation, celle de Limber (Kitching, Alsing et al., 2016), (valable pour un multipôle  $l < 100$ , soit une échelle angulaire supérieure à 2 degrés) permet d'écrire :

$$C_{ij}^{\gamma\gamma}(\ell) \simeq \frac{c}{H_0} \int dz \frac{W_i^\gamma(z)W_j^\gamma(z)}{E(z)r^2(z)} P_{\delta\delta} \left[ \frac{\ell + 1/2}{r(z)}, z \right], \quad (4.2)$$

où  $i$  et  $j$  sont les paires de bins de redshift,  $E(z)$  l'expression sans dimension du taux de Hubble,  $r(z)$  la distance comobile,  $P_{\delta\delta}$  est le spectre de puissance matière évalué à  $k = k_l(z) \sim (l + 1/2)/r(z)$  à cause de l'approximation de Limber. Nous définissons aussi la fonction "fenêtre" ("Weight function") comme ceci :

$$W_i^\gamma(z) = \frac{3}{2} \frac{H_0}{c} \Omega_{m,0} (1+z) \tilde{r}(z) \int_z^{z_{\max}} dz' n_i(z') \left[ 1 - \frac{\tilde{r}(z)}{\tilde{r}(z')} \right] \quad (4.3)$$

$$= \frac{3}{2} \frac{H_0}{c} \Omega_{m,0} (1+z) \tilde{r}(z) \widetilde{W}_i(z). \quad (4.4)$$

où le facteur  $\tilde{r}(z)$  est la distance comobile divisé par le rayon de Hubble  $c/H_0$  ( $\tilde{r}(z) = r(z)/(c/H_0)$ ) pour souligner la forte dépendance entre le taux d'expansion réduit  $h$  et la fonction "fenêtre"  $W_i^\gamma(z)$  avec le facteur  $(H_0/c)$ . Nous introduisons aussi la fonction "fenêtre" réduite  $\widetilde{W}_i(z)$  qui simplifie par la suite les calculs.

### Spectre de puissance d'alignement intrinsèque

Durant la formation des galaxies, des processus de marée et d'autres processus engendrent une forme préférentielle et corrélée de formes galactiques, mis en évidence par la fonction de cisaillement à 2 points. Cet effet, nommé **Intrinsic Alignment (IA)** peut être considéré comme une erreur astrophysique systématique et est difficilement supprimable des données brutes. La contribution de l'IA peut être vue en considérant la somme de 4 termes de fonctions de corrélations à 2 points :

$$C_{ij}^{\epsilon\epsilon}(\ell) = C_{ij}^{\gamma\gamma}(\ell) + C_{ij}^{I\gamma}(\ell) + C_{ij}^{\gamma I}(\ell) + C_{ij}^{II}(\ell), \quad (4.5)$$

où le premier terme  $C_{ij}^{\gamma\gamma}(\ell)$  est défini en (4.2). Le second terme  $C_{ij}^{I\gamma}(\ell)$  représente la corrélation entre l'arrière-plan de cisaillement et l'avant-plan de l'alignement intrinsèque. Le troisième terme  $C_{ij}^{\gamma I}(\ell)$  évalue la corrélation entre l'avant-plan de cisaillement et l'arrière plan d'ellipticité. Enfin, le dernier terme est nul car un cisaillement d'avant-plan ne peut pas être corrélé avec un arrière-plan d'ellipticité à part si les redshifts des galaxies sont mal ciblés.

### Redshifts photométriques et densité

La densité  $n_i(z)$  de galaxies observées dans le bin  $i$  a pour distribution, avec une estimation de redshift photométrique :

$$n_i(z) = \frac{\int_{z_i^-}^{z_i^+} dz_p n(z) p_{\text{ph}}(z_p|z)}{\int_{z_{\min}}^{z_{\max}} \int_{z_i^-}^{z_i^+} dz dz_p n(z) p_{\text{ph}}(z_p|z)}, \quad (4.6)$$

avec  $(z_i^-, z_i^+)$  qui sont les valeurs des cotés du bin de redshift  $i$ .

À titre d'exemple, dans le cas du Weak Lensing sur LSST, j'ai pris pour le cas Y10 (10 années d'observation) la formule suivante issue du fit de (LSST Dark Energy Science Collaboration et al., 2018) avec les valeurs ( $z_0 = 0.11$  et  $\alpha_0 = 0.68$ ) :

$$n(z) = z^2 \exp \left[ - \left( \frac{z}{z_0} \right)^{\alpha_0} \right] \quad (4.7)$$

Nous utiliserons aussi les spécifications définissant un intervalle équidistant de 10 bins pour le cas Y10 :

$$z_i = \{0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95, 1.05, 1.15\}, \quad (4.8)$$

et les valeurs des 2 cotés de chaque bin centré :

$$z_i^- = \{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.10\}, \quad (4.9)$$

$$z_i^+ = \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2\}, \quad (4.10)$$

En ce qui concerne la vraie distribution sous-jacente  $n(z)$  à l'expression de (4.6), celle-ci est convoluée avec la distribution de probabilité  $p_{ph}(z_p|z)$  décrivant la probabilité qu'une galaxie avec un redshift  $z$  soit mesurée avec un redshift  $z_p$  (Kitching, Amara et al., 2009). Une paramétrisation pratique pour cette quantité est donnée par :

$$p_{ph}(z_p|z) = \frac{1 - f_{out}}{\sqrt{2\pi}\sigma_b(1+z)} \exp \left\{ -\frac{1}{2} \left[ \frac{z - c_b z_p - z_b}{\sigma_b(1+z)} \right]^2 \right\} + \frac{f_{out}}{\sqrt{2\pi}\sigma_o(1+z)} \exp \left\{ -\frac{1}{2} \left[ \frac{z - c_o z_p - z_o}{\sigma_o(1+z)} \right]^2 \right\}, \quad (4.11)$$

Cette expression permet d'inclure un terme multiplicatif et un terme additif pour une proportion  $(1 - f_{out})$  de sources dont le redshift est mesuré relativement bien et une fraction  $f_{out}$  avec des valeurs incohérentes. Nous pouvons modifier les paramètres de cette équation mais dans le code de notre survey GC<sub>phot</sub>, ils sont fixés et font partie des spécifications de LSST. D'autres analyses cosmologiques permettraient d'effectuer un calibrage plus précis sur ces paramètres.

### Le shot noise

La partie non-corrélée de l'ellipticité intrinsèque (c'est-à-dire non lentillée) agit comme un terme de bruit dans le spectre de puissance. On fixe ce bruit à une valeur non nulle pour les spectres auto-corrélés (à l'intérieur d'un même bin) et à 0 pour des cross-corrélation (plusieurs valeurs par bin), car les ellipticités des galaxies à différents redshifts ne sont pas corrélées.

### Spectre de puissance de cisaillement cosmique observé

Finalement, dans le cas flat (courbure nulle) et avec l'approximation de Limber (Kitching, Alsing et al., 2016), nous pouvons exprimer le spectre de puissance tomographique de cisaillement :

$$C_{ij}^{\epsilon\epsilon}(\ell) = C_{ij}^{\gamma\gamma}(\ell) + C_{ij}^{\Pi}(\ell) + C_{ij}^{\Gamma}(\ell) + N_{ij}^{\epsilon}(\ell) \quad (4.12)$$

Voir aussi l'Appendice B pour le développement des différents termes.

### 4.2.2 Formalisme de Fisher appliqué au Weak Lensing

L'élément  $(\alpha, \beta)$  de la matrice de Fisher s'exprime sous la forme :

$$F_{\alpha\beta} = f_{sky} \sum_{\ell=\ell_{min}}^{\ell_{max}} \left( \ell + \frac{1}{2} \right) \Delta\ell \frac{\partial C_{ij}^{\epsilon\epsilon}(\ell)}{\partial p_{\alpha}} \left[ \mathbf{C}^{\epsilon\epsilon}(\ell)^{-1} \right]_{jm} \frac{\partial C_{mn}^{\epsilon\epsilon}(\ell)}{\partial p_{\beta}} \left[ \mathbf{C}^{\epsilon\epsilon}(\ell)^{-1} \right]_{ni} \quad (4.13)$$

avec  $\mathbf{C}^{\epsilon\epsilon}(\ell)$  la matrice tomographique et  $[\mathbf{C}^{\epsilon\epsilon}(\ell)^{-1}]_{jm}$  l'élément  $(j, m)$  de l'inverse de la matrice tomographique au multipôle  $l$ . La sommation sur tous les indices répétés est implicite.

Pour construire la matrice de Fisher, nous avons besoin, en plus de la matrice tomographique, de ses dérivées par rapport à chacun des paramètres cosmologiques. Contrairement au Galaxy Clustering de DESI, nous n'effectuons pas de projection et pouvons directement remplir la matrice avec les paramètres désirés.

### 4.3 Formalisme de Fisher appliqué au Galaxy clustering photométrique

Pour le survey Galaxy Clustering photométrique, nous utilisons la distribution de galaxies  $n(z)$  de (4.7) convoluée avec la probabilité  $p_{ph}$  exprimée en (4.11).

Nous notons d'abord le contraste de densité au redshift  $i$  par le terme  $\delta_g^i$ , la source du cisaillement du bin tomographique  $j$  comme  $\gamma^j$  et la densité de galaxies au bin tomographique  $i$  pour le survey GC photométrique comme  $n_i(z)$ . La densité angulaire pour ce bin de redshift est donc :

$$\bar{n}_i = \int dz n_i(z) \quad (4.14)$$

où  $n_i$  est donnée par (4.6). Nous pouvons maintenant définir la fonction de "fenêtre radiale" pour le GC clustering :

$$W_i^{\delta_g}(k, z) = b_i(k, z) \frac{n_i(z)}{\bar{n}_i} H(z), \quad (4.15)$$

avec  $H(z)$  la constante de Hubble-Lemaître en  $\text{km.s}^{-1}.\text{Mpc}^{-1}$  et  $b_i(k, z)$  le biais de la galaxie dans le bin tomographique  $i$ . Pour cette dernière quantité, nous supposons une valeur constante pour chaque redshift sans dépendance sur l'échelle  $k$ . Cette valeur pour chaque redshift est considérée comme un paramètre de nuisance et la quantité de référence est donnée par la formule :

$$b_i(z) = \sqrt{1 + \bar{z}}, \quad (4.16)$$

avec  $\bar{z}$  le redshift moyen de chaque bin. Finalement, d'après la même formule que celle utilisée pour le GC spectroscopique, on obtient l'expression du spectre de puissance du GC photométrique :

$$C_{ij}^{\delta_g \delta_g}(\ell) = \int \frac{dz}{H(z)r^2(z)} \mathcal{W}_i^{\delta_g}(z) \mathcal{W}_j^{\delta_g}(z) P_{\delta\delta} \left( \frac{\ell + 1/2}{r(z)}, z \right) \quad (4.17)$$

avec le terme  $\mathcal{W}_i^{\delta_g}$  exprimé en (4.15).

### 4.4 Formalisme de Fisher appliqué à la combinaison du Weak Lensing avec le Galaxy Clustering photométrique

Le codage du croisement de données entre la sonde  $\text{GC}_{\text{phot}}$  et WL dépasse le cadre et les objectifs de cette étude. Cependant, nous pouvons résumer la combinaison des résultats issus de ces 2 sondes en disant que les éléments de la matrice de Fisher sont obtenus selon une formule similaire à celle vue en (4.13) mais cette fois-ci il nous faut faire intervenir une matrice de covariance construite d'une manière différente car dépendante à la fois du WL et de  $\text{GC}_{\text{phot}}$ .

### 4.5 Synthèse finale des FoM de LSST

Le module XSAF du code `Totally SAF` permet de calculer, avec les spécifications que j'ai introduites, les matrices de Fisher pour le Galaxy clustering photométrique, le Weak Lensing et les matrices de cross-corrélations entre ces 2 sondes. Voici ci-dessous l'ensemble des FoM obtenues pour  $\text{GC}_{\text{phot}}$  et WL dans les cas Y1, Y10, flat et non-flat :

**TABLE 4.2 – FoM (Figure of Merit) pour les 8 cas possibles de configuration.**

	WL FLAT Y1	WL NON-FLAT Y1	WL FLAT Y10	WL NON-FLAT Y10
FoM	0.5676	0.1678	7.3559	2.1331
	$\text{GC}_{\text{phot}}$ FLAT Y1	$\text{GC}_{\text{phot}}$ NON-FLAT Y1	$\text{GC}_{\text{phot}}$ FLAT Y10	$\text{GC}_{\text{phot}}$ NON-FLAT Y10
FoM	56.8890	33.6088	659.4959	353.4001

Comme avec le survey DESI, nous voyons que les performances sont meilleures dans le cas flat que dans le cas non-flat ; le paramètre supplémentaire de courbure, que l'on estime par la valeur de l'énergie noire, rend l'estimation de l'ensemble moins précise : on rajoute ainsi de l'incertitude. Il faut aussi remarquer que le temps d'observation sur 10 années d'observation (Y10) améliore les contraintes par rapport au cas d'une seule année (Y1), ce qui est logique car les données sont plus nombreuses et donc de l'information supplémentaire est apportée.

## 4.6 Synthèse finale des matrices de Fisher

Si on assimile les sondes  $GC_{\text{phot}}$  et WL à 2 variables aléatoires indépendantes, alors la matrice de Fisher totale du survey LSST est la somme des matrices de Fisher  $GC_{\text{phot}}$  ( $F_{GCp}$ ) et WL ( $F_{WL}$ ) :

$$F_{\text{tot}} = F_{GCp} + F_{WL} \quad (4.18)$$

Les 2 matrices à ajouter ayant des tailles différentes, il s'agit de construire une matrice totale tenant compte à la fois des paramètres cosmologiques, des différents biais pour les 2 sondes, des 3 paramètres d'alignement pour le WL et des cas Univers "flat" et "non-flat".

## 4.7 Comparaison des résultats pour LSST

Voici ci-dessous les FoM obtenues à partir desquelles nous pouvons faire une comparaison entre le cas simple, où l'on ajoute les matrices de Fisher issues du  $GC_{\text{phot}}$  avec celles venant du Weak Lensing, et le cas "cross-corrélé" où l'on calcule directement, via la matrice de Fisher avec croisement de données, les contraintes provenant de la matrice de Fisher générée par le code XSAF. Dans ce cas-là, on ne considère plus les sondes  $GC_{\text{phot}}$  et WL comme étant indépendantes.

**TABLE 4.3 – Synthèse des FoM pour ( $GC_{\text{phot}} + WL$ ) versus ( $GC_{\text{phot}} + WL + XC$ ) dans les cas (Y1, Y10) et (Flat, Non-Flat).**

Synthèse des FoM avec ( $GC_{\text{phot}} + WL$ ) comparé à ( $GC_{\text{phot}} + WL + XC$ )				
	WL FLAT-Y1	WL NON-FLAT-Y1	$GC_{\text{phot}}$ FLAT-Y1	$GC_{\text{phot}}$ NON-FLAT-Y1
FoM	0.57	0.17	56.89	33.61
	WL FLAT-Y10	WL NON-FLAT-Y10	$GC_{\text{phot}}$ FLAT-Y10	$GC_{\text{phot}}$ NON-FLAT-Y10
FoM	7.36	2.13	659.50	353.40
	( $GC_{\text{phot}} + WL$ ) FLAT-Y1	( $GC_{\text{phot}} + WL$ ) NON-FLAT-Y1	( $GC_{\text{phot}} + WL$ ) FLAT-Y10	( $GC_{\text{phot}} + WL$ ) NON-FLAT-Y10
FoM	64.44	44.37	737.60	556.93
	( $GC_{\text{phot}} + WL + XC$ ) FLAT-Y1	( $GC_{\text{phot}} + WL + XC$ ) NON-FLAT-Y1	( $GC_{\text{phot}} + WL + XC$ ) FLAT-Y10	( $GC_{\text{phot}} + WL + XC$ ) NON-FLAT-Y10
FoM	122.24	61.88	1599.25	741.39

Comme cela était attendu, pour une même période d'observation (1 an ou 10 ans) et une même sonde, les FoM pour le cas flat sont plus élevées que celles du cas non-flat. Ce résultat avait déjà été mis en évidence pour le survey DESI : l'explication de cette différence est la même, c'est-à-dire que le fait de rajouter un paramètre supplémentaire à estimer diminue la précision des estimations de tous les autres paramètres.

Nous voyons aussi que la durée d'observation influe directement sur les contraintes : une collecte de 10 ans de données (Y10) donnera une précision supérieure sur les contraintes par rapport à une collecte sur 1 an (Y1). Là aussi, de manière intuitive, le volume de données étant plus important, nous disposons d'échantillons supplémentaires, et donc d'un gain d'informations.

De plus, nous pouvons noter qu'il y a un gain non négligeable sur la sonde  $GC_{\text{phot}}$  version "matrices sommées", c'est-à-dire en incluant le WL, par rapport à la version unique de  $GC_{\text{phot}}$  : ceci est valable pour une même durée d'observation ou une même géométrie (flat, non-flat).

Enfin, on remarque que la "cross-corrélation" apporte beaucoup plus de précision sur les contraintes comparée à une simple somme du  $GC_{\text{phot}}$  et du WL : en effet, la cross-corrélation permet de croiser des données différentes pour un même bin de redshift alors que la somme uniquement de la matrice de Fisher sur le Galaxy clustering photométrique et celle du Weak Lensing ne tient pas compte de cette combinaison de données, qui apporte, au sens de la théorie de l'information de Shannon, une incertitude moindre sur les paramètres et donc un gain en contraintes.

## 4.8 Synthèse des résultats DESI-LSST

<b>DESI(S/W/M)</b>	<b>Flat</b>	<b>Non-Flat</b>
GCsp(S)	114.73	51.66
GCsp(W)	145.17	62.09
GCsp(M)	155.93	64.47
<b>LSST(Y1)+DESI(M)</b>	<b>Flat</b>	<b>Non-Flat</b>
GCph	56.89	33.61
WL	0.57	0.17
GCph+WL	64.44	44.37
GCph+WL+XC	122.24	61.88
GCsp(M)+GCph+WL	446.31	197.86
GCsp(M)+GCph+WL+XC	528.29	205.37
<b>LSST(Y10)+DESI(M)</b>	<b>Flat</b>	<b>Non-Flat</b>
GCph	659.50	353.40
WL	7.36	2.13
GCph+WL	737.60	556.93
GCph+WL+XC	1599.25	741.39
GCsp(M)+GCph+WL	1302.12	835.62
GCsp(M)+GCph+WL+XC	2051.82	1000.60

TABLE 4.4 – *Synthèse des surveys DESI et LSST avec toutes les sondes et différentes combinaisons (Flat/Non-Flat, (S = population unique - fusion de colonnes, W = population unique - moyenne pondérée, M = Multi populations) et Y1/Y10)*

### 4.9 Comparaison avec le survey EUCLID

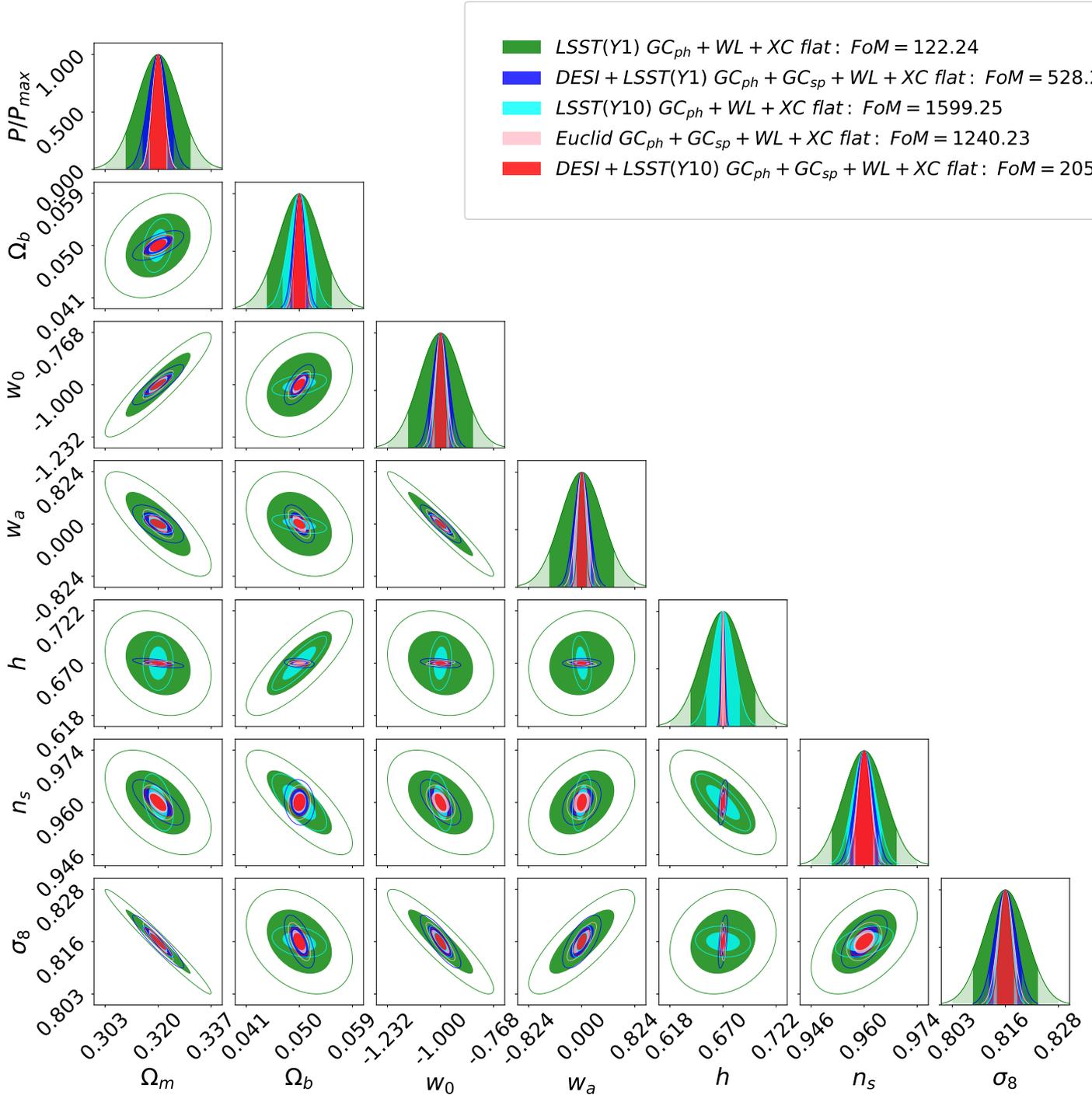


FIGURE 4.1 – Contours représentant dans le cas flat les distributions posterior pour la comparaison entre notre étude et le survey EUCLID

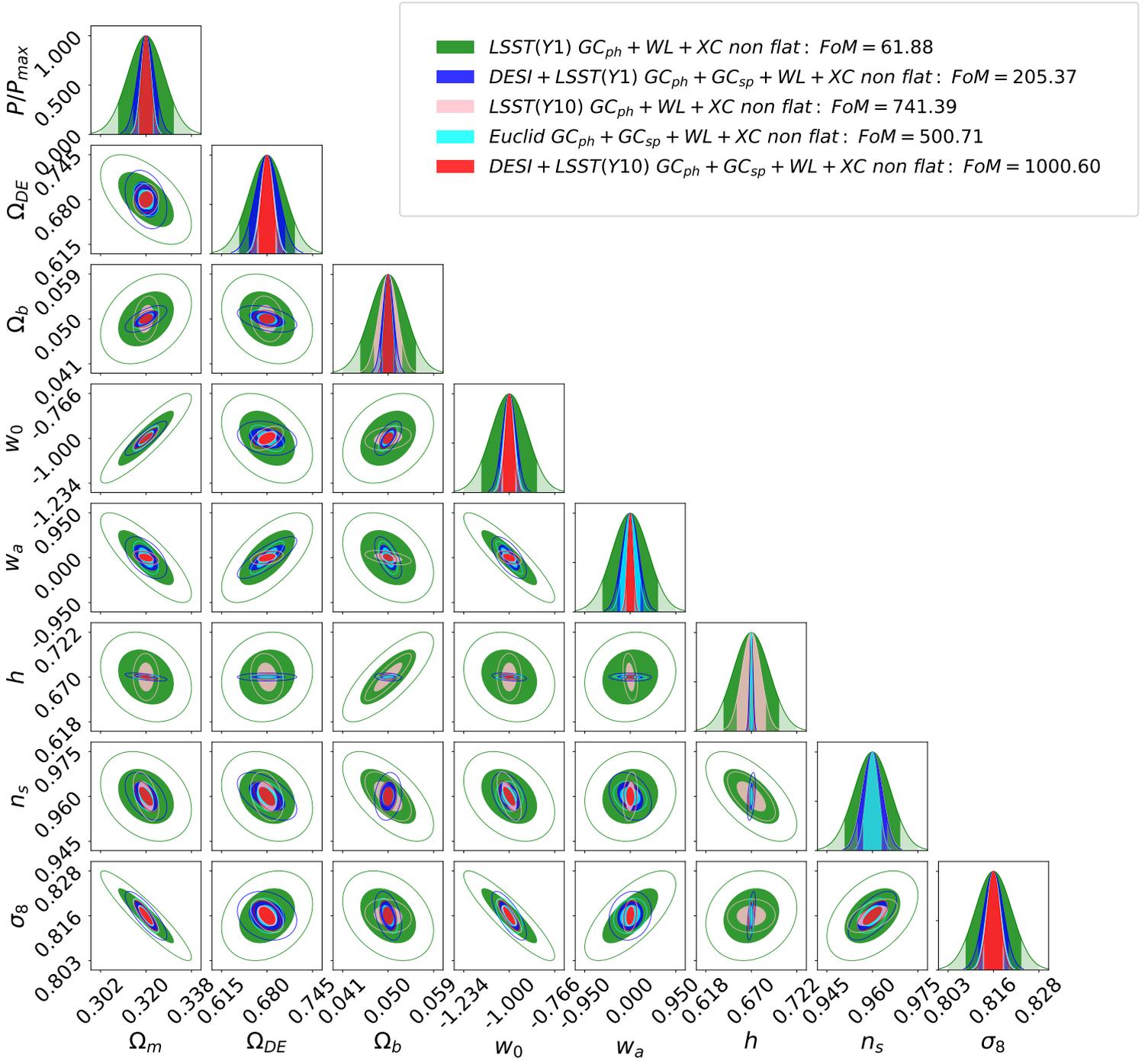


FIGURE 4.2 – Contours représentant dans le cas non-flat les distributions posterior pour la comparaison entre notre étude et le survey EUCLID

# Conclusion

Notre étude nous a permis d'obtenir, grâce au forecast effectué pour les 2 surveys DESI et LSST, des contraintes en accord avec ce qui était attendu au niveau des ordres de grandeur. Pour le survey DESI et sa sonde Galaxy clustering spectroscopique, nous avons utilisé 2 méthodes différentes afin d'essayer d'améliorer les résultats initiaux obtenus avec les données originales.

La première, étant assimilée à une moyenne pondérée des biais sur les densités de galaxies, est donc un traitement "single population" puisque nous ne faisons pas du croisement de données à proprement parler : la double dépendance en redshift et en type de population, évoquée ci-après dans la méthode "multi-traceurs", n'est pas prise en considération. Elle produit une légère augmentation du niveau des contraintes de départ (+27% en flat et +20% en non-flat).

La deuxième, faisant intervenir une matrice de covariance des observables, apporte un gain dans les cas d'Univers plat/non-plat mais il est relativement minime par rapport à ce que l'on aurait pu espérer. Cette méthode est qualifiée de "multi-traceurs" ou encore "multi-populations". En effet, nous tenons compte, pour les valeurs de biais entrecroisés, à la fois d'une dépendance sur les redshifts et d'une dépendance sur les types de populations considérées. Le gain que l'on obtient en faisant cette cross-corrélation n'est pas très élevé car d'une part, nous ne disposons pas de beaucoup d'échantillons et d'autre part, le faible nombre de valeurs de biais enchevêtrées pour un même bin de redshift est susceptible de produire moins d'informations au sens de Fisher. Malgré ces conditions assez défavorables sur le plan statistique, la cross-corrélation permet de réduire les incertitudes sur l'estimation des paramètres (+36% en flat et +25% en non flat).

Concernant le survey LSST, les contraintes produites nous indique tout d'abord que la sonde Galaxy Clustering photométrique est beaucoup plus performante que celle du Weak Lensing. Ceci est directement lié à aux valeurs des densités de galaxies : il a été démontré, dans d'autres études, qu'à faible densité, la sonde  $GC_{ph}$  produisait des FoM et contraintes meilleures qu'avec la sonde WL. Pour des densités plus importantes, à partir d'un certain seuil, la tendance s'inverse et c'est la sonde WL qui devient supérieure au niveau de la précision des estimations alors que pour la sonde  $GC_{ph}$ , la FoM sature.

En combinant les 3 sondes des relevés DESI et LSST, c'est-à-dire ( $GC_{sp}, GC_{ph}, WL$ ), nous augmentons encore plus la précision des évaluations. Le bénéfice est particulièrement significatif avec le rajout de la partie cross-corrélée, en atteignant par exemple une FoM  $> 2000$  pour le cas "LSST(10 ans d'observation) cross-corrélé + DESI(Multi-populations)".

Enfin, nous avons pu confirmer l'importance des cross-corrélations entre les sondes  $GC_{ph}$  et WL : la FoM est multipliée par un facteur compris environ entre 1.5 (non-flat) et 2 (flat) dès lors qu'elles sont prises en compte.

## Perspectives avec le survey EUCLID

Nous avons aussi comparé cette synthèse DESI-LSST avec les forecasts réalisés sur la mission EUCLID dont le lancement est prévu pour 2022. Nous nous sommes rendus compte que, indépendamment de la géométrie plate ou non plate de l'Univers, le relevé LSST avec 10 ans d'observation sera le plus performant des 3 surveys, suivi par le relevé EUCLID avec ses 6 ans et demi d'observation puis par le relevé DESI. Cependant, ces 3 surveys seront complémentaires et apporteront, une fois tous leurs résultats conjugués, des contraintes plus fortes sur les paramètres caractérisant l'énergie noire.

# Bibliographie

- Albrecht, A. et al. (sept. 2006). “Report of the Dark Energy Task Force”. In : *Astronomy and Astrophysics Advisory Committee*. arXiv : [astro-ph/0609591](#).
- Ballinger, W. E., J. A. Peacock et A. F. Heavens (oct. 1996). “Measuring the cosmological constant with redshift surveys”. In : *mnras* 282, p. 877. DOI : [10.1093/mnras/282.3.877](#). eprint : [astro-ph/9605017](#).
- Chevallier, M. et D. Polarski (sept. 2000). “Accelerating universes with scaling Dark Matter”. In : *Int.J.Mod.Phys. D10 :213-224*. arXiv : [arXiv:gr-qc/0009008](#).
- Desjacques, V., D. Jeong et F. Schmidt (nov. 2016). “Large-Scale Galaxy Bias”. In : *ArXiv e-prints*. arXiv : [1611.09787](#).
- Eisenstein, D. J., H.-J. Seo et M. White (août 2007). “On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter”. In : *apj* 664, p. 660-674. DOI : [10.1086/518755](#). eprint : [astro-ph/0604361](#).
- Euclid Collaboration et al. (2018). “Forecast validation for the Euclid satellite”. In : *Article in prep*.
- Hamilton, A. J. S. (1998). “Linear Redshift Distortions : a Review”. In : *Astrophysics and Space Science Library* 231. Sous la dir. de D. Hamilton, p. 185. DOI : [10.1007/978-94-011-4960-0\\_17](#). eprint : [astro-ph/9708102](#).
- Kaiser, N. (juil. 1987). “Clustering in real space and in redshift space”. In : *mnras* 227, p. 1-21. DOI : [10.1093/mnras/227.1.1](#).
- Kitching, T. D., J. Alsing et al. (nov. 2016). “The Limits of Cosmic Shear”. In : *ArXiv e-prints*. arXiv : [1611.04954](#).
- Kitching, T. D., A. Amara et al. (nov. 2009). “Cosmological systematics beyond nuisance parameters : form-filling functions”. In : *mnras* 399, p. 2107-2128. DOI : [10.1111/j.1365-2966.2009.15408.x](#). arXiv : [0812.1966 \[astro-ph\]](#).
- Lewis PA, Challinor A., Lasenby A. (2000). “Efficient Computation of CMB anisotropies in closed FRW models”. In : *ApJ J.538*, p. 473-476. arXiv : [astro-ph/9911177](#).
- Linder, E. V. (nov. 2003). “Mapping the Dark Energy Equation of State”. In : *Phys. Rev. Lett.* 90, 091301. arXiv : [astro-ph/0311403](#).
- LSST Dark Energy Science Collaboration et al. (2018). “The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document”. In : arXiv : [1809.01669v1](#).
- Peebles, P.J.E (1980). “The Large Scale Structure of the Universe”. In : *Princeton University Press*.

- Perlmutter, S. et al. (1998). “Measurements of  $\Omega$  and  $\Lambda$  from 42 high-redshift supernovae”. In : *ApJ*. arXiv : [astro-ph/9812133](#).
- Robert Scherrer, J. (mai 2015). “Mapping the Chevallier-Polarski-Linder parametrization onto Physical Dark Energy Models”. In : *Phys. Rev. D* 92, 043001. arXiv : [arXiv:1505.05781](#).
- Seo, H.-J. et D. J. Eisenstein (déc. 2003). “Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys”. In : *ApJ* 598, p. 720-740. arXiv : [astro-ph/0307460](#).
- Taylor, P. L. et al. (avr. 2018). “Testing the Cosmic Shear Spatially-Flat Universe Approximation with GLaSS”. In : *ArXiv e-prints*. arXiv : [1804.03668](#).
- Tegmark, Max (nov. 1997). “Measuring Cosmological Parameters with Galaxy Surveys”. In : *Phys. Rev. Lett.* 79, p. 3806-3809. arXiv : [astro-ph/9706198](#).
- Wang, Bin et al. (2007). “Interacting Dark Energy and Dark Matter : observational Constraints from Cosmological Parameters”. In : *Nucl. Phys.* B778, p. 69-84. DOI : [10.1016/j.nuclphysb.2007.04.037](#). arXiv : [astro-ph/0607126](#).
- Wang, Y., C.-H. Chuang et C. M. Hirata (avr. 2013). “Towards more realistic forecasting of dark energy constraints from galaxy redshift surveys”. In : *mnras* 430, p. 2446-2453. DOI : [10.1093/mnras/stt068](#). arXiv : [1211.0532](#).
- Wang, Y., W. Percival et al. (déc. 2010). “Designing a space-based galaxy redshift survey to probe dark energy”. In : *mnras* 409, p. 737-749. DOI : [10.1111/j.1365-2966.2010.17335.x](#). eprint : [1006.3517](#).
- White, M., Y. Song et W. Percival (2008). “Forecasting Cosmological Constraints from Redshift Surveys”. In : *ApJ*. arXiv : [arXiv:0810.1518](#).

# Galaxy clustering - Démonstration de l'expression du spectre de puissance observable

Les effets que nous avons énumérés dans la section 3.6, à savoir : **1) le biais des galaxies**, **2) anisotropies dues aux distorsions spatiales en redshift** dont nous avons parlé dans la section 3.5 (RSD effect), **3) le bruit de fond**, **4) l'incertitude sur les redshifts**, et **5) les effets de projection** dont l'effet **Alcock-Pazcynski (AP)**, interviennent tous dans le calcul du modèle du spectre de puissance (nous allons adapter son expression dans le cadre de notre forecast pour la sonde Galaxy Clustering). Nous ferons la démonstration dans le régime linéaire puis non-linéaire.

**1) Biais effectif des galaxies** : la notion de biais relie la variation de densité de matière noire à la distribution de matière à un niveau plus global. Pour plus de précisions, je renvoie le lecteur à la référence suivante (Desjacques, Jeong et Schmidt, 2016). L'expression du spectre de puissance fonction du biais des galaxies  $P_{g,\text{lin}}$  est une simple relation linéaire avec le spectre de puissance linéaire  $P_m(k, z)$  qui s'écrit :

$$P_{g,\text{lin}}(k; z) = b^2(z)P_m(k; z) \quad (\text{A.1})$$

Le lien entre les 2 spectres de puissance est simplement le carré du biais effectif de l'échantillon  $b(z)$ .

**2) Anisotropies dues aux RSD** : la mesure d'un redshift d'une galaxie est composé d'une partie non-cosmologique qui s'appelle la vitesse péculière le long de la ligne de visée, c'est-à-dire une vitesse locale supplémentaire qui doit s'ajouter ou se retirer à la vitesse classique issue de la loi de Hubble-Lemaître. L'expression du redshift  $z_{\text{obs}}$  observé est :

$$1 + z_{\text{obs}} = (1 + z) (1 + v_{\parallel}/c) , \quad (\text{A.2})$$

où  $v_{\parallel}$  est la vitesse péculière. Si l'on considère tous les redshifts d'origine purement cosmologique, on remarque un type particulier d'anisotropies sur le spectre de puissance observé : c'est l'effet RSD. Comme la distribution des vitesses péculières et celle de la densité sont liées, cet effet nous renseigne sur leur influence mutuelle. Dans le régime linéaire de cet effet RSD, la relation du spectre de puissance réel s'exprime de cette façon (Kaiser, 1987) :

$$P_{\text{zs},\text{lin}}(k, \mu; z) = [b(z)\sigma_8(z) + f(z)\sigma_8(z)\mu^2]^2 \frac{P_m(k; z)}{\sigma_8^2(z)} , \quad (\text{A.3})$$

avec  $\sigma_8(z)$  défini en 3.3 mais avec en plus une dépendance en  $z$ . Cette relation (A.3) illustre bien le fait que les 2 paramètres principaux influant sur le spectre de puissance sont les termes  $b(z)\sigma_8(z)$  and  $f(z)\sigma_8(z)$  que nous traitons comme des paramètres libres.

**3) Incertitude sur les redshifts et bruit de fond** : l'effet des incertitudes sur le redshift et du shot noise résiduel peut être modélisé ainsi :

$$P_{\text{zerr},\text{lin}}(k, \mu; z) = P_{\text{zs},\text{lin}}(k; z)F_z(k, \mu; z) + P_s(z) , \quad (\text{A.4})$$

avec  $P_s(z)$  un offset indépendant d'échelle dû à l'imperfection de la suppression du shot noise qui est fixé à  $P_s(z) = 0$  pour tous les redshifts de notre modèle fiducial, mais qui peut varier en fonction du redshift dans le calcul de notre forecast. Concernant le facteur  $F_z(k, \mu; z)$ , il est égal à :

$$F_z(k, \mu; z) = e^{-k^2 \mu^2 \sigma_r^2(z)}, \quad (\text{A.5})$$

Ce facteur explique la dégradation du champ de densité des galaxies le long de la ligne de visée  $k_{\parallel} = k\mu$  en raison d'erreurs possibles sur la mesure des redshifts. L'erreur  $\sigma_r$  se propage selon la distance comobile (Y. Wang, Chuang et Hirata, 2013) :

$$\sigma_r(z) = \frac{\partial r}{\partial z} \sigma_z(z) = \frac{c}{H(z)} (1+z) \sigma_0. \quad (\text{A.6})$$

#### 4) Les effets de projection AP

La mesure du spectre de puissance des galaxies requiert un modèle cosmologique de référence pour pouvoir convertir les redshifts en distances. En supposant un modèle non valide, cela provoquera une obligation de remettre à la bonne échelle le vecteur d'onde  $\vec{k}$ , dans les directions parallèles et perpendiculaires à la ligne de visée :

$$k_{\perp} = \frac{k_{\perp, \text{ref}}}{q_{\perp}} \quad \text{and} \quad k_{\parallel} = \frac{k_{\parallel, \text{ref}}}{q_{\parallel}}, \quad (\text{A.7})$$

où les coefficients  $q_{\perp}$  et  $q_{\parallel}$  sont les quotients entre la distance angulaire réelle et de référence, et entre le paramètre de Hubble réel et de référence :

$$q_{\perp}(z) = \frac{D_A(z)}{D_{A, \text{ref}}(z)} \quad \text{and} \quad q_{\parallel}(z) = \frac{H_{\text{ref}}(z)}{H(z)} \quad (\text{A.8})$$

Nous enlevons le label  $z$  pour une notation plus légère et prenons les variables  $(k_{\text{ref}}, \mu_{\text{ref}})$  pour désigner les valeurs de référence. Nous en déduisons :

$$\begin{aligned} k(k_{\text{ref}}, \mu_{\text{ref}}) &= \frac{k_{\text{ref}}}{q_{\perp}} \left[ 1 + \mu_{\text{ref}}^2 \left( \frac{q_{\perp}^2}{q_{\parallel}^2} - 1 \right) \right]^{1/2}, \quad \text{et} \\ \mu(\mu_{\text{ref}}) &= \mu_{\text{ref}} \frac{q_{\perp}}{q_{\parallel}} \left[ 1 + \mu_{\text{ref}}^2 \left( \frac{q_{\perp}^2}{q_{\parallel}^2} - 1 \right) \right]^{-1/2}. \end{aligned} \quad (\text{A.9})$$

Ces relations peuvent être utilisées pour analyser l'influence du modèle cosmologique choisi sur le spectre de puissance observé en cartographiant les valeurs  $(k_{\text{ref}}, \mu_{\text{ref}})$  pour trouver celles qui sont les plus pertinentes (Ballinger, Peacock et Heavens, 1996) :

$$P_{\text{proj, lin}}(k_{\text{ref}}, \mu_{\text{ref}}; z) = \frac{1}{q_{\perp}^2 q_{\parallel}} P_{\text{zs, lin}}(k(k_{\text{ref}}, \mu_{\text{ref}}), \mu(\mu_{\text{ref}}); z), \quad (\text{A.10})$$

#### Modèle linéaire complet du spectre de puissance observé :

D'après tout ce qui précède, nous avons l'expression complète du spectre de puissance observé, linéaire et anisotropique :

$$P_{\text{obs}}(k_{\text{ref}}, \mu_{\text{ref}}; z) = \frac{1}{q_{\perp}^2 q_{\parallel}} (b\sigma_8(z) + f\sigma_8(z)\mu^2)^2 \frac{P_m(k; z)}{\sigma_8^2(z)} F_z(k, \mu; z) + P_s(z), \quad (\text{A.11})$$

avec  $k = k(k_{\text{ref}}, \mu_{\text{ref}})$  et  $\mu = \mu(\mu_{\text{ref}})$ .

## Échelle non linéaire pour le Galaxy clustering :

À une époque relativement jeune en terme de temps cosmique, et aux petites échelles, la distribution de matière est affectée par l'évolution non linéaire des fluctuations de densité, ce qui modifie la forme du spectre de puissance au delà des prédictions des modèles de perturbation linéaire (Y. Wang, Chuang et Hirata, 2013).

Le premier effet de non-linéarité est que l'on doit rajouter un facteur de type Lorentzien à la description RSD : ce facteur rend compte de l'effet Finger-of-God (FoG) (Hamilton, 1998). Deuxièmement, nous devons tenir compte de l'amortissement dans les oscillations acoustiques des baryons (BAO). Nous considérons 2 composantes d'amortissement : radiale et transverse. Le spectre de puissance associé à cette atténuation est qualifié de "de-wiggled". On peut écrire alors :

$$P_{zs}(k, \mu; z) = \left\{ \frac{1}{1 + [f(z)k\mu \sigma_p(z)]^2} \right\} [b(z)\sigma_8(z) + f(z)\sigma_8(z)\mu^2]^2 \frac{P_{dw}(k, \mu; z)}{\sigma_8^2(z)}, \quad (\text{A.12})$$

avec  $\sigma_p(z)$  l'écart type des paramètres de nuisance le long de la ligne de visée et le spectre  $P_{dw}(k, \mu; z)$  ("de-wiggled") qui tient compte de l'amortissement des BAO (Y. Wang, Chuang et Hirata, 2013) :

$$P_{dw}(k, \mu; z) = P_m(k; z) e^{-g_\mu k^2} + P_{nw}(k, \mu; z) \left(1 - e^{-g_\mu k^2}\right). \quad (\text{A.13})$$

$P_m$  représente toujours le spectre linéaire alors que  $P_{nw}$  est un "no-wiggle" spectre, c'est-à-dire ne tenant pas compte des caractéristiques des BAO. Concernant la fonction  $g_\mu(k, \mu, z)$ , elle est égale à :

$$g_\mu(k, \mu, z) = \sigma_v^2(z) \left\{ 1 - \mu^2 + \mu^2 [1 + f(z)]^2 \right\} \quad (\text{A.14})$$

Cette relation traduit l'amortissement non-linéaire des BAO par l'effet FoG (Eisenstein, Seo et White, 2007). Le paramètre  $\sigma_v(z)$  contrôle l'intensité de cet amortissement. En principe,  $\sigma_v(z)$  et  $\sigma_p(z)$  représentent la même quantité physique : la dispersion des vitesses dans un modèle linéaire, qui peuvent être calculées par :

$$\sigma_v^2(z) = \sigma_p^2(z) = \frac{1}{3} \int \frac{1}{(2\pi)^3} \frac{P_{\theta\theta}}{k^2} d^3k = \frac{1}{6\pi^2} \int P_m(k, z) dk \quad (\text{A.15})$$

En pratique, cette relation nous permet de définir les valeurs "fiduciales" (de référence) de  $\sigma_v(z)$  et  $\sigma_p(z)$  ; on prend :

$$\sigma_v(z) = \sigma_v(z_{\text{mean}}) \frac{D(z)}{D(z_{\text{mean}})}, \quad (\text{A.16})$$

$$\sigma_p(z) = \sigma_p(z_{\text{mean}}) \frac{D(z)}{D(z_{\text{mean}})}, \quad (\text{A.17})$$

avec  $z_{\text{mean}}$  le redshift central du survey et  $D(z)$  le facteur de croissance ("growth rate"). Nous considérons les 2 quantités ci-dessus comme paramètres libres, ce qui permet de mieux évaluer au cours du temps la croissance de l'impact non-linéaire sur l'évolution des petites et grandes structures.

Nous pouvons ainsi dire que  $\sigma_v(z)$  et  $\sigma_p(z)$  sont étroitement liés aux vitesses de dispersion des galaxies entre elles.

## Modèle non-linéaire complet du spectre de puissance observé :

On retrouve finalement l'équation (3.7) qui est implémentée dans notre code de Galaxy Clustering pour le survey DESI :

$$P_{\text{obs}}(k, \mu; z) = \frac{1}{q_\perp^2 q_\parallel} \left\{ \frac{[b\sigma_8(z) + f\sigma_8(z)\mu^2]^2}{1 + [f(z)k\mu\sigma_p(z)]^2} \right\} \frac{P_{dw}(k, \mu; z)}{\sigma_8^2(z)} F_z(k, \mu; z) + P_s(z) \quad (\text{A.18})$$

# Formalisme de Fisher appliqué au Weak Lensing

Dans le cas flat (courbure nulle) et avec l'approximation de Limber (Kitching, Alsing et al., 2016), le spectre de puissance tomographique de cisaillement s'écrit :

$$C_{ij}^{\epsilon\epsilon}(\ell) = C_{ij}^{\gamma\gamma}(\ell) + C_{ij}^{\text{II}}(\ell) + C_{ij}^{\text{I}\gamma}(\ell) + N_{ij}^{\epsilon}(\ell) \quad (\text{B.1})$$

qui s'exprime aussi en développant les différents termes :

$$\begin{aligned} C_{ij}^{\epsilon\epsilon}(\ell) &= \frac{c}{H_0} \int dz \frac{W_i^{\gamma}(z)W_j^{\gamma}(z)}{E(z)r^2(z)} P_{\delta\delta} \left[ \frac{\ell + 1/2}{r(z)}, z \right] \\ &+ \frac{c}{H_0} \int dz \frac{W_i^{\gamma}(z)W_j^{\text{IA}}(z) + W_i^{\text{IA}}(z)W_j^{\gamma}(z)}{E(z)r^2(z)} P_{\delta\text{I}} \left[ \frac{\ell + 1/2}{r(z)}, z \right] \\ &+ \frac{c}{H_0} \int dz \frac{W_i^{\text{IA}}(z)W_j^{\text{IA}}(z)}{E(z)r^2(z)} P_{\text{II}} \left[ \frac{\ell + 1/2}{r(z)}, z \right] \\ &+ N_{ij}^{\epsilon}(\ell), \end{aligned} \quad (\text{B.2})$$

Les 3 premières lignes traduisent respectivement la contribution du cisaillement théorique cosmologique, le cisaillement avec "l'alignement intrinsèque" (shear-IA) et les corrélations entre "alignements intrinsèques" (IA-IA correlations).

Nous rappelons au lecteur que les indices  $i$  et  $j$  correspondent aux labels des redshifts dans lesquels la distribution du redshift des sources a été divisé. Chaque terme de (B.2) a la même structure générale : tout d'abord la présence d'une intégrale sur le redshift du produit de fonctions "fenêtre"  $W_{\mu}^{\nu}$  vues en (4.3) (dépendant seulement de l'évolution cosmique) et d'un spectre de puissance  $P_{\alpha\beta}$  (caractérisant la croissance des structures cosmiques).

L'intégration se fait en principe jusqu'au redshift de l'horizon ( $z \simeq 1100$ ) mais en pratique, nous sommes beaucoup plus contraints car le survey est limité par  $z \leq z_{max}$ . De la même manière, la limite inférieure de l'intégration ne peut pas être égale à 0 et tout comme la borne supérieure, les performances et la sensibilité de l'instrument impacteront cette borne inférieure. Par conséquent, d'après les spécifications énoncées en 4.1, l'intervalle d'intégration pour le survey LSST sera  $(z_{min}, z_{max}) = (0.2, 1.2)$ .