

Etude statistique des corrélations dans la distribution des galaxies - application aux Oscillations Baryoniques Acoustiques

Antoine Labatie

Laboratoire CosmoStat, Service d'Astrophysique, CEA Saclay

Directeurs de thèse:

Jean-Luc Starck

Marc Lachièze-Rey

Introduction

- Le modèle standard Λ CDM de la cosmologie permet de bien comprendre toutes les observations
- Mais il fait intervenir une forme inconnue d'énergie, **l'énergie sombre**
 - Comprendre sa nature est un objectif majeur de la cosmologie moderne
- On peut notamment étudier l'énergie sombre avec la propriété d'étalon standard des Oscillations Baryoniques Acoustiques (BAOs) dans les relevés de galaxies
- Comme il s'agit d'un effet statistique très faible, il faut faire attention aux effets systématiques
 - **Mon travail s'inscrit dans l'étude de certains effets statistiques systématiques**

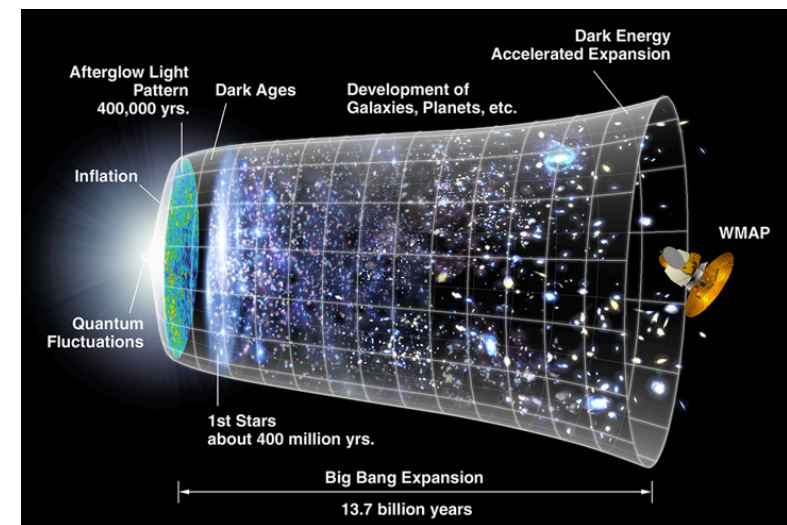
Plan

- Modèle Λ CDM et BAOs
- La fonction de corrélation ξ
- Relevés de galaxies
- Incertitude sur l'estimation de ξ
- Modélisation statistique de ξ avec des simulations
- Application à la détection des BAOs
- Application aux contraintes cosmologiques

I) Modèle Λ CDM et BAOs

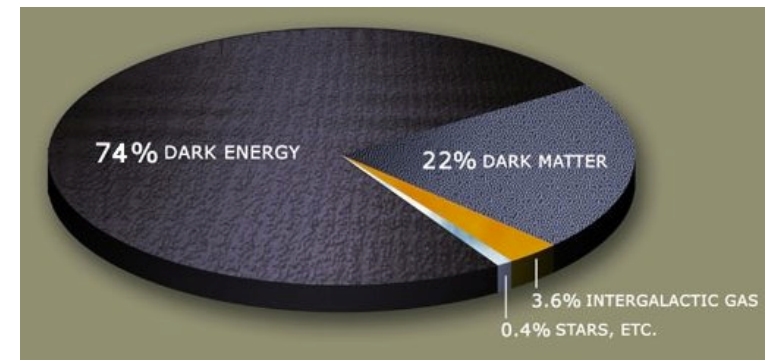
Le modèle Λ CDM

- Repose sur le modèle du Big Bang chaud:
 - Plasma primordial extrêmement dense et chaud
 - Période d'inflation avec expansion exponentielle
 - Nucléosynthèse primordiale (formation des premiers noyaux)
 - Recombinaison (formation des premiers atomes) $\approx 380\,000$ ans après Big Bang. Le découplage des photons et baryons provoque l'émission du fond diffus cosmologique (CMB)
 - Evolution gravitationnelle, formation des grandes structures



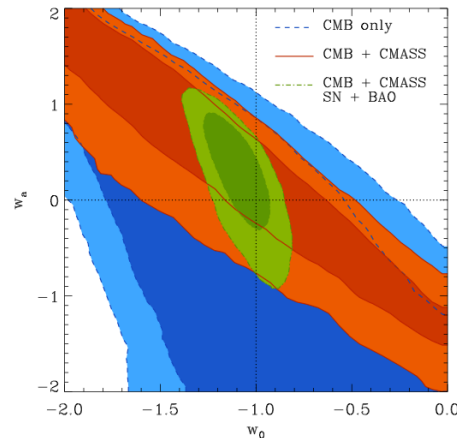
Le modèle Λ CDM

- Modèle cosmologique standard apparu à la fin des années 90 en accord avec les $\neq$ observations (CMB, grandes structures, abondance éléments légers, Supernovae Ia, ...)
- L'univers Λ CDM est constitué de:
 - $\approx 74\%$ d'énergie noire sous forme d'une constante cosmologique Λ
 - $\approx 22\%$ de matière noire froide
 - $\approx 4\%$ de matière baryonique



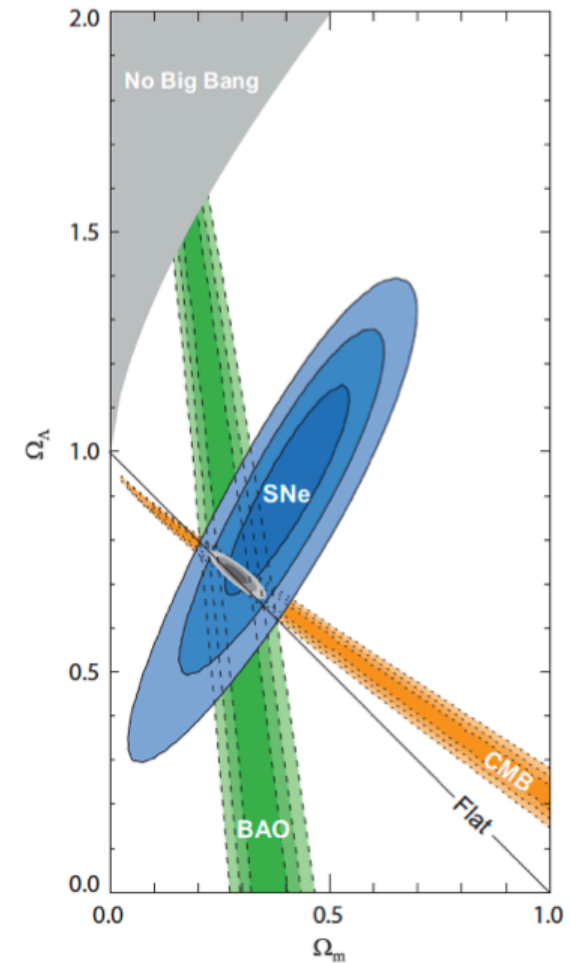
Objectifs de la cosmologie moderne

- Contrainte des paramètres cosmologiques ($\Omega_\Lambda, \Omega_m, \Omega_b, n_s, \sigma_8, H_0$)
- Etude de l'énergie sombre Ω_{DE}, w_{DE} ($w_\Lambda = -1$)



$$w(a) = w_0 + w_a(1 - a) \quad a = 1/(1 + z)$$

Sanchez et al. (2012)

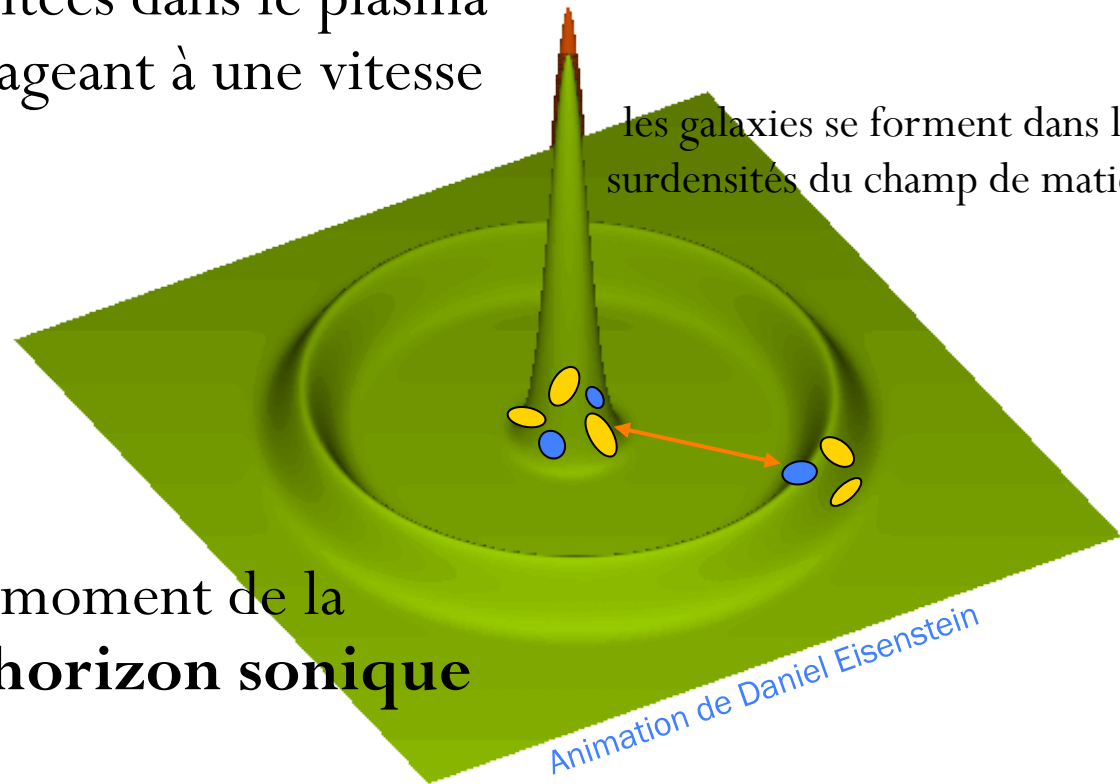


Kowalski et al. (2008)

Oscillations Baryoniques Acoustiques

- Ondes sonores excitées dans le plasma primordial se propageant à une vitesse $\approx c/\sqrt{3}$

les galaxies se forment dans les surdensités du champ de matière

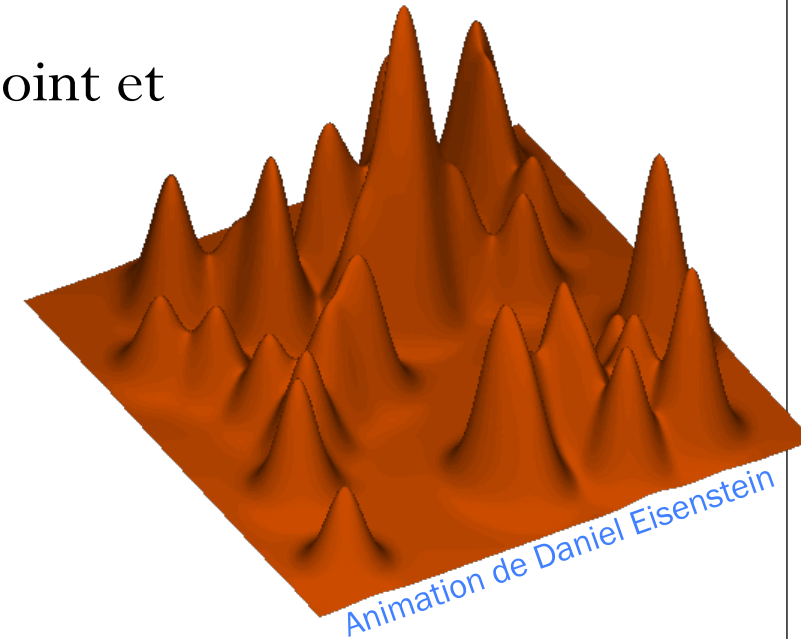
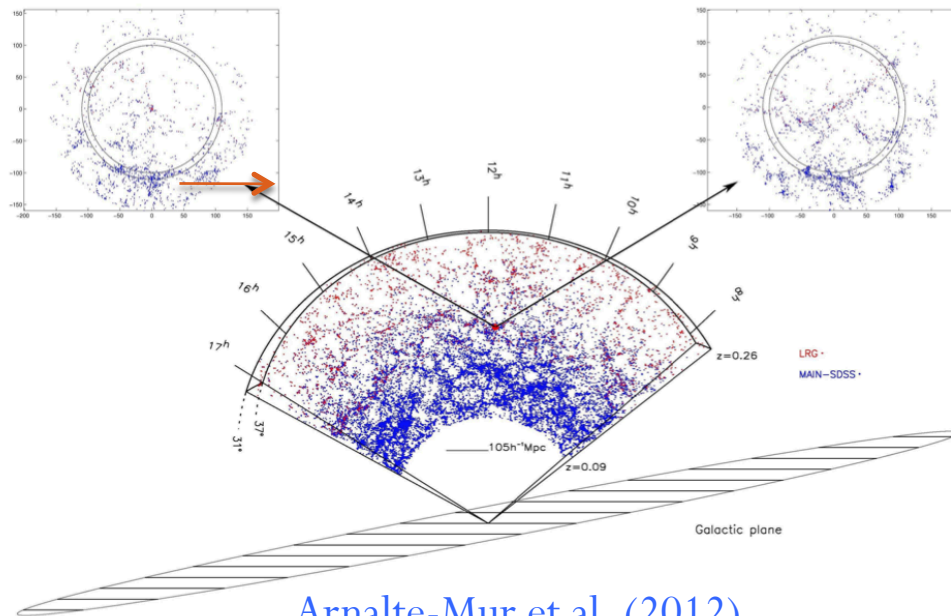


- L'onde s'arrête au moment de la recombinaison à **l'horizon sonore**
 $r_s \approx 150 \text{ Mpc}$

Animation de Daniel Eisenstein

Oscillations Baryoniques Acoustiques

- Ces ondes sont générées en tout point et se superposent
- Effet statistique de seulement 1%



Animation de Daniel Eisenstein

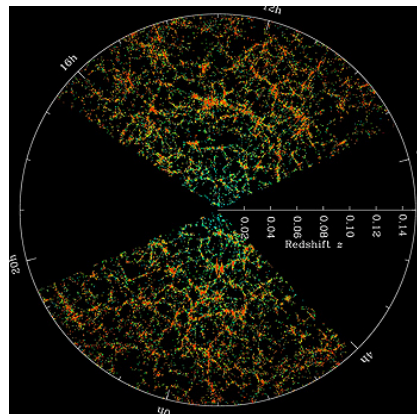


Evolution gravitationnelle des grandes structures

BAOs = étalon standard

- Les relevés de galaxies sont des relevés en redshift $\longrightarrow$ on doit **supposer une cosmologie fiducielle** pour obtenir un volume 3D

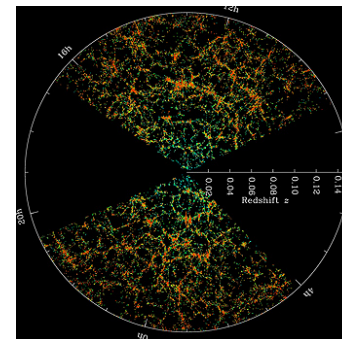
cosmologie fiducielle
correcte



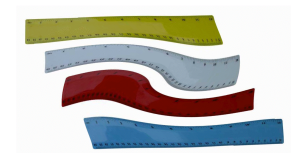
$$D_C = D_H \int_0^z \frac{dz'}{E(z')}$$

cosmologie fiducielle
incorrecte

$\longrightarrow$ distortions



- Les BAOs fournissent un étalon standard (taille réelle connue)
 $\longrightarrow$ ils montrent comment la cosmologie fiducielle est incorrecte



II) La fonction de corrélation

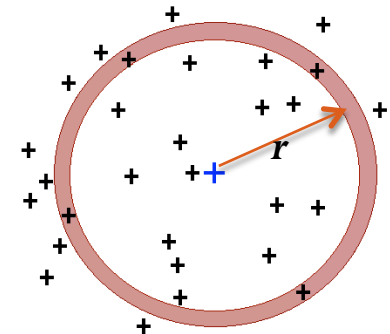
La fonction de corrélation

- Probabilité de trouver une galaxie dans volume δV :

$$\delta P = \bar{n} \delta V$$

- Probabilité de trouver une paire de galaxies dans des volumes δV_1 et δV_2 séparés de $\mathbf{r}$:

$$\delta P_{12} = \bar{n}^2 [1 + \xi(\mathbf{r})] \delta V_1 \delta V_2$$



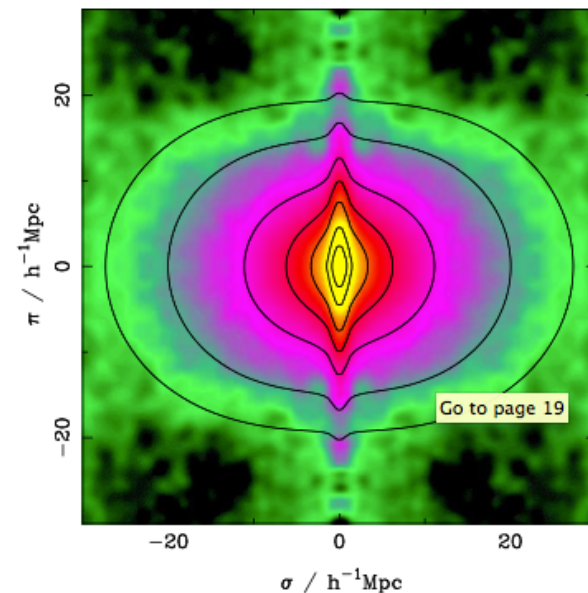
excès de galaxies à une distance $\mathbf{r}$

- $\xi(\mathbf{r})$ donne l'excès de probabilité de trouver une paire à une séparation $\mathbf{r}$ comparé à une distribution aléatoire

La fonction de corrélation Λ CDM

- Fonction de corrélation de la matière linéaire $\xi_L(r)$ dépendant de $\Omega_m h^2$, $\Omega_b h^2$, n_s (à un facteur d'amplitude près)
- Evolution gravitationnelle non linéaire de la matière
- Biais des galaxies par rapport à la matière
- Distortions de redshift:

$$\xi_0(s) = \frac{1}{4\pi} \int_{\Omega} \xi(\mathbf{s}) d^2 o_{\mathbf{s}}$$
$$\xi_0(s) = \left(1 + \frac{2\beta}{3} + \frac{\beta^2}{5}\right) \xi(r)$$



Peacock et al. (2001)

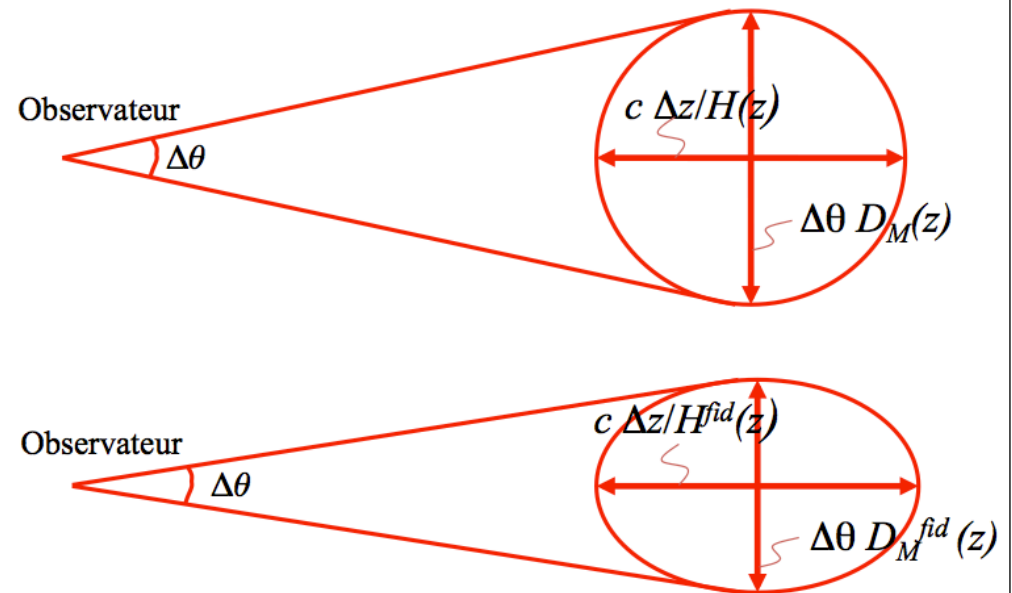
La fonction de corrélation Λ CDM

- Effet de la cosmologie fiducielle:

$$\mathbb{E} \left[\hat{\xi}(\pi, \sigma) \right] = \xi \left(\frac{H^{fid}(z)}{H(z)} \pi, \frac{D_M(z)}{D_M^{fid}(z)} \sigma \right)$$

π : séparation dans la ligne de visée

σ : séparation dans la direction transverse



La fonction de corrélation Λ CDM

- On définit l'échelle de dilatation:

$$D_V(z) = \left[(1+z)^2 D_M(z)^2 \frac{cz}{H(z)} \right]^{1/3}$$

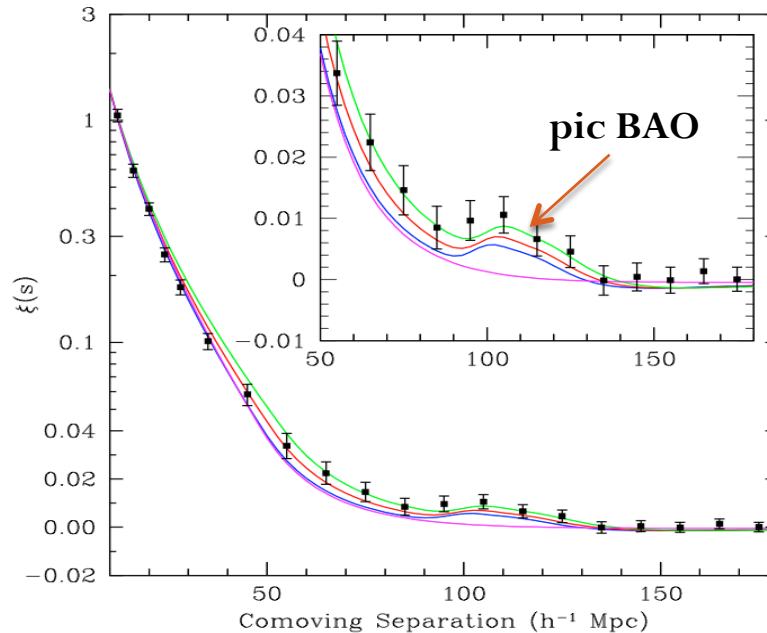
- L'effet de la cosmologie fiducielle peut être approximé comme une simple dilatation du monopole (Eisenstein et al. (2005), Padmanabhan and White (2008))

$$\alpha = \frac{D_V(\langle z \rangle)}{D_V^{fid}(\langle z \rangle)}$$

$$\mathbb{E} \left[\hat{\xi}_0(s) \right] = \xi_0(\alpha s)$$

- α est directement relié aux paramètres d'énergie sombre Ω_{DE} , w_{DE}
 - En effet l'énergie sombre influence fortement les quantités $D_M(z)$ et $H(z)$ aux redshifts faibles

La fonction de corrélation Λ CDM



$\xi(r)$ pour $\Omega_m h^2 = 0.12, 0.13, 0.14$

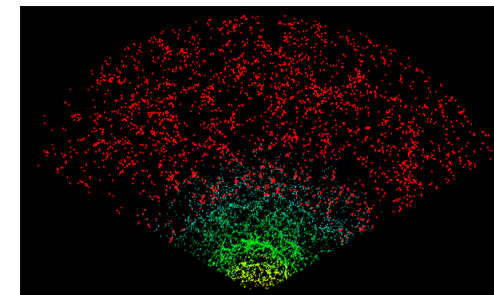
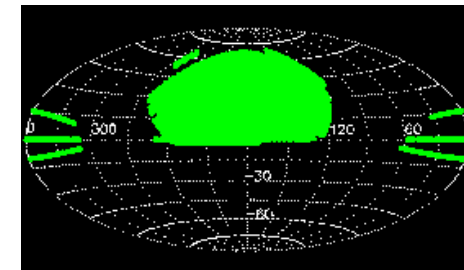
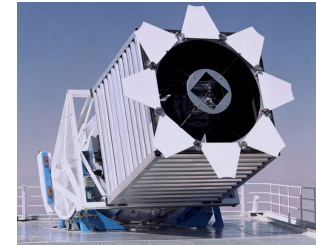
Eisenstein et al. (2005)

- La détection des BAOs à l'échelle attendue supporte le modèle Λ CDM (prédiction théorique de Peebles en 1970)
- Les BAOs contraignent α (et donc l'énergie sombre) avec la propriété d'étalon standard

III) Relevés de galaxies

Le relevé SDSS DR7

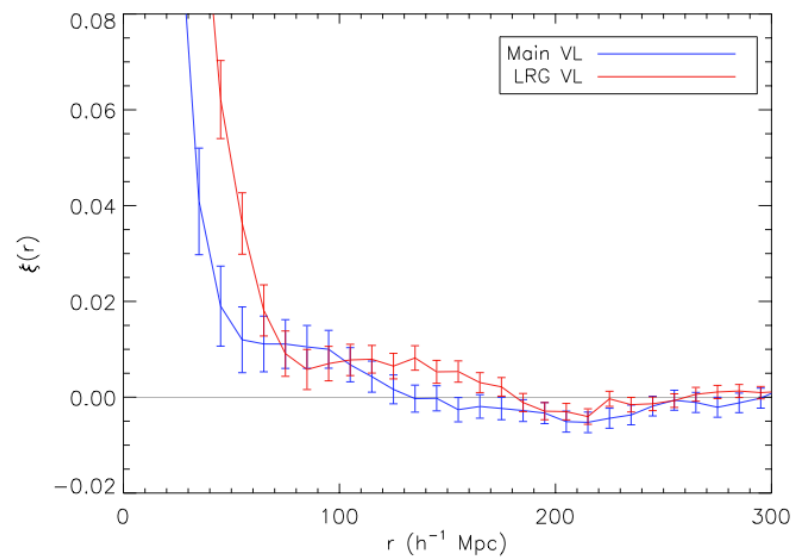
- Programme de 8 ans avec un télescope de 2.5m au point Apache (Nouveau Mexique)
- Cartographié 7500 deg² du ciel
- Spectre de 930 000 galaxies
- 1 échantillon limité en magnitude apparente jusqu'à une distance $D \approx 600 h^{-1} \text{Mpc}$ (Main)
- 1 échantillon de densité presque constante (volume-limited) jusqu'à une distance $D \approx 1150 h^{-1} \text{Mpc}$ (LRG)



Catalogues étudiés

- Catalogues SDSS DR7 volume-limited:

	Limites de magnitude	Limites de <i>redshift</i>	Surface (deg ²)	N_g	V (h^{-1} Gpc) ³	$\bar{n}$ (h^{-1} Mpc) ⁻³
Main VL	$M_r < -21$	$0.059 < z < 0.168$	4,809	67,189	61.2×10^{-3}	1.098×10^{-3}
LRG VL	$-23.544 < M_g < -21.644$	$0.14 < z < 0.42$	4,809	34,347	0.79	4.345×10^{-5}

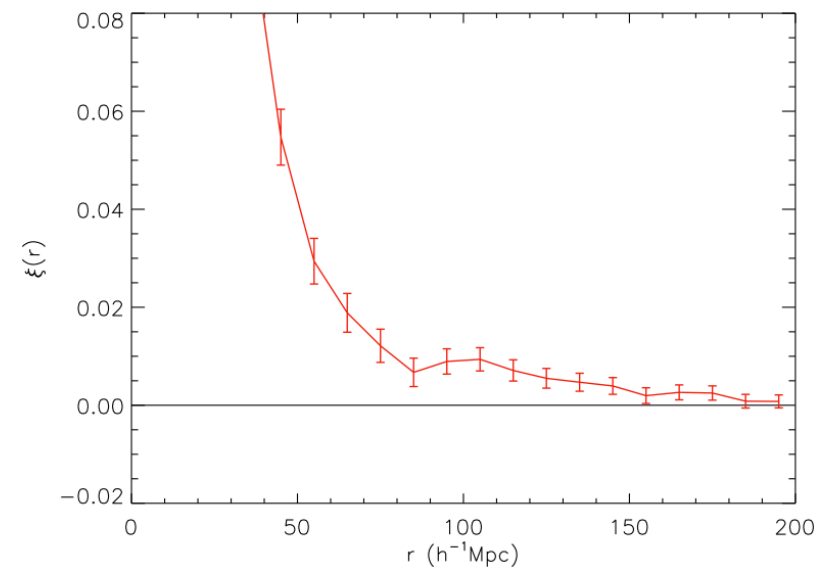


Catalogues étudiés

- Catalogue SDSS DR7-Full (LRG)

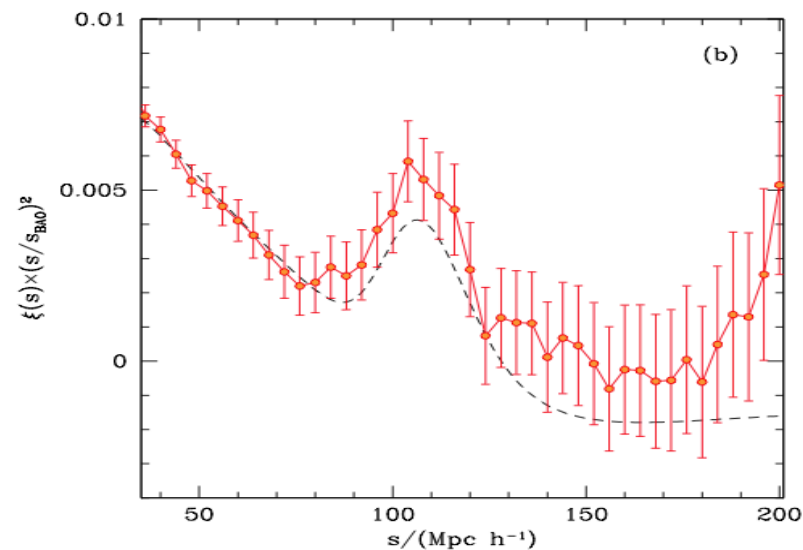
	Limites de magnitude	Limites de <i>redshift</i>	Surface (deg ²)	N_g	V (h^{-1} Gpc) ³	$\bar{n}$ (h^{-1} Mpc) ⁻³
DR7-Full	$-23.2 < M_g < -21.2$	$0.16 < z < 0.47$	6,918	105,831	1.58	6.70×10^{-5}

- Pic BAO plus large que prévu:
 - Faible détection des BAOs
 - Déplacement du maximum de vraisemblance vers $\alpha < 1$



Catalogues étudiés

- Catalogue BOSS DR9 CMASS
 - Relevé SDSS III (2009-2014), extension du relevé SDSS II
 - Données DR9 publiques en août 2012
 - Fort signal BAO dans échantillon DR9 CMASS



Sanchez et al. (2012)

IV) Incertitude sur l'estimation de ξ

-Quelle est l'importance du biais dû à la contrainte intégrale sur les estimateurs de la fonction de corrélation ?

Labatie et al. (2012 a)

Estimation de la fonction de corrélation

- Estimation soumise à des *incertitudes*:
 - Seulement un volume fini observé (variance cosmique)
 - Seulement un nombre fini de galaxies (shot noise)
- On utilise des catalogues aléatoires pour calculer l'excès de paires à une distance r
- Différents estimateurs:
 - Peebles-Hauser (1974)
$$\hat{\xi}_{PH}(r) = \frac{N_{RR}}{N_{DD}} \frac{DD(r)}{RR(r)} - 1$$
 - Davis-Peebles (1983)
$$\hat{\xi}_{DP}(r) = \frac{N_{DR}}{N_{DD}} \frac{DD(r)}{DR(r)} - 1$$
 - Hamilton (1993)
$$\hat{\xi}_{HAM}(r) = \frac{N_{DR}^2}{N_{DD}N_{RR}} \frac{DD(r)RR(r)}{[DR(r)]^2} - 1$$
 - Landy-Szalay (1993)
$$\hat{\xi}_{LS}(r) = 1 + \frac{N_{RR}}{N_{DD}} \frac{DD(r)}{RR(r)} - 2 \frac{N_{RR}}{N_{DR}} \frac{DR(r)}{RR(r)}$$

La contrainte intégrale

- La densité moyenne est estimée à partir du catalogue lui-même:

$$\hat{\xi}_{PH}(r) = \frac{N_{RR}}{N_{DD}} \frac{DD(r)}{RR(r)} - 1$$

*estimé par # total de paires au lieu de
l'espérance du # total de paires*

$$\delta P_{12} = \bar{n}^2 [1 + \xi(\mathbf{r})] \delta V_1 \delta V_2$$

*estimée par la densité moyenne du catalogue au
lieu de la vraie densité moyenne*

- Cela cause la contrainte intégrale sur les estimateurs:

$$\int \hat{\xi}(\mathbf{r}) d^3\mathbf{r} \approx 0$$

La contrainte intégrale

- Toutefois la vraie fonction de corrélation $\xi(\mathbf{r})$ ne vérifie par nécessairement cette contrainte:

$$\int \xi(\mathbf{r}) d^3\mathbf{r} \neq 0$$

- Cela impose un biais aux estimateurs:

$$\mathbb{E} \left[\int \hat{\xi}(\mathbf{r}) d^3\mathbf{r} \right] = \int \mathbb{E} [\hat{\xi}(\mathbf{r})] d^3\mathbf{r} = 0 \neq \int \xi(\mathbf{r}) d^3\mathbf{r}$$

La contrainte intégrale

- En réalité il y a une fonction de pondération dans l'intégrale de ξ
 - On peut montrer que l'estimateur de Peebles-Hauser vérifie presque exactement la contrainte intégrale suivante:

$$f(r) = \frac{1}{V^2} \int_V d^3\mathbf{x} \int_V d^3\mathbf{y} \mathbb{1}_{|\mathbf{y}-\mathbf{x}| \in [r \pm dr/2]}$$

$$\int_0^{r_{max}} f(r) \hat{\xi}_{PH}(r) dr = 0$$

- Les autres estimateurs vérifient la contrainte seulement de façon approximative:

$$g(r) = \frac{1}{N_D} \sum_{\mathbf{d}_j} \frac{1}{V} \int_V d^3\mathbf{y} \mathbb{1}_{|\mathbf{y}-\mathbf{d}_j| \in [r \pm dr/2]}$$

$$\approx f(r)$$

$$\int_0^{r_{max}} f(r) \hat{\xi}_{DP}(r) dr \approx 0$$

$$\int_0^{r_{max}} f(r) \hat{\xi}_H(r) dr \approx 0$$

$$\int_0^{r_{max}} f(r) \hat{\xi}_{LS}(r) dr \approx 0$$

La contrainte intégrale

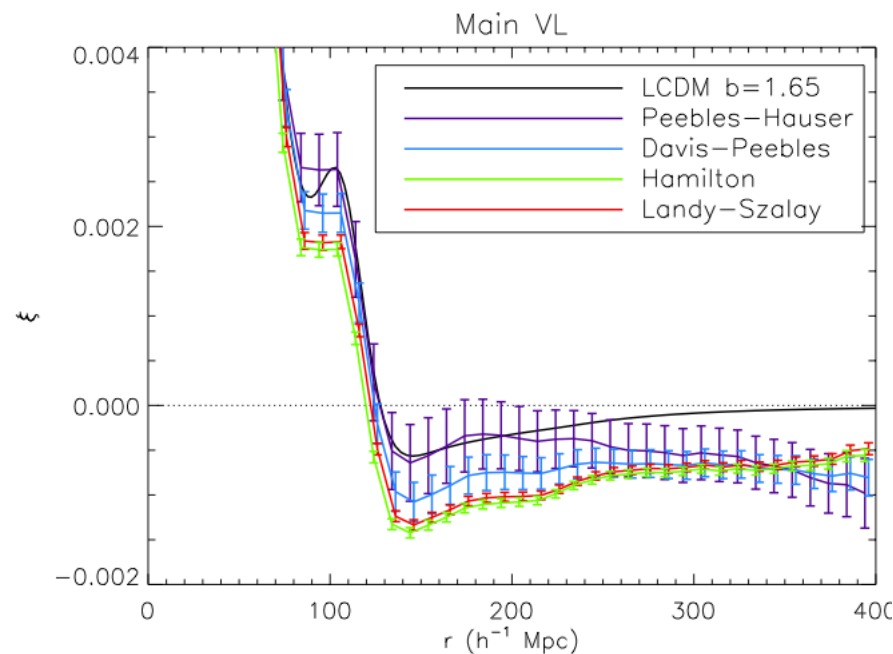
- Dans un univers fractal, la contrainte intégrale a un effet important quelque soit la taille du catalogue:
 - Idée soutenue par un groupe de chercheurs ([Labini et al.](#)), selon lesquels la contrainte intégrale rend l'estimation biaisée
- Si on suppose l'homogénéité à grande échelle (le principe cosmologique), la contrainte intégrale n'a qu'un effet pour des catalogues de taille limitée
 - **On veut ici vérifier l'importance de cet effet**

Méthode d'analyse

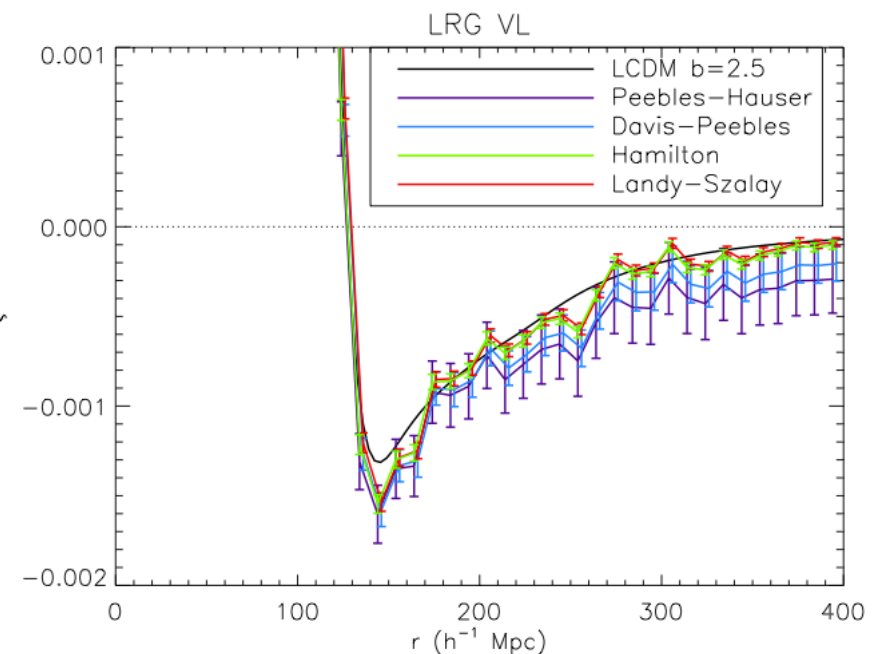
- On suppose un modèle standard Λ CDM et on génère des **simulations log-normales des catalogues SDSS Main VL et LRG VL**:
- On quantifie l'effet de la contrainte intégrale: Y a-t-il un biais dans l'estimation de ξ pour les catalogues de galaxies actuels?
- On compare les différents estimateurs de ξ en termes de biais et de variance

Biais de la contrainte intégrale pour les $\neq$ estimateurs

- Moyennes empiriques des $\neq$ estimateurs pour 2000 simulations log-normales du Main VL et du LRG VL:

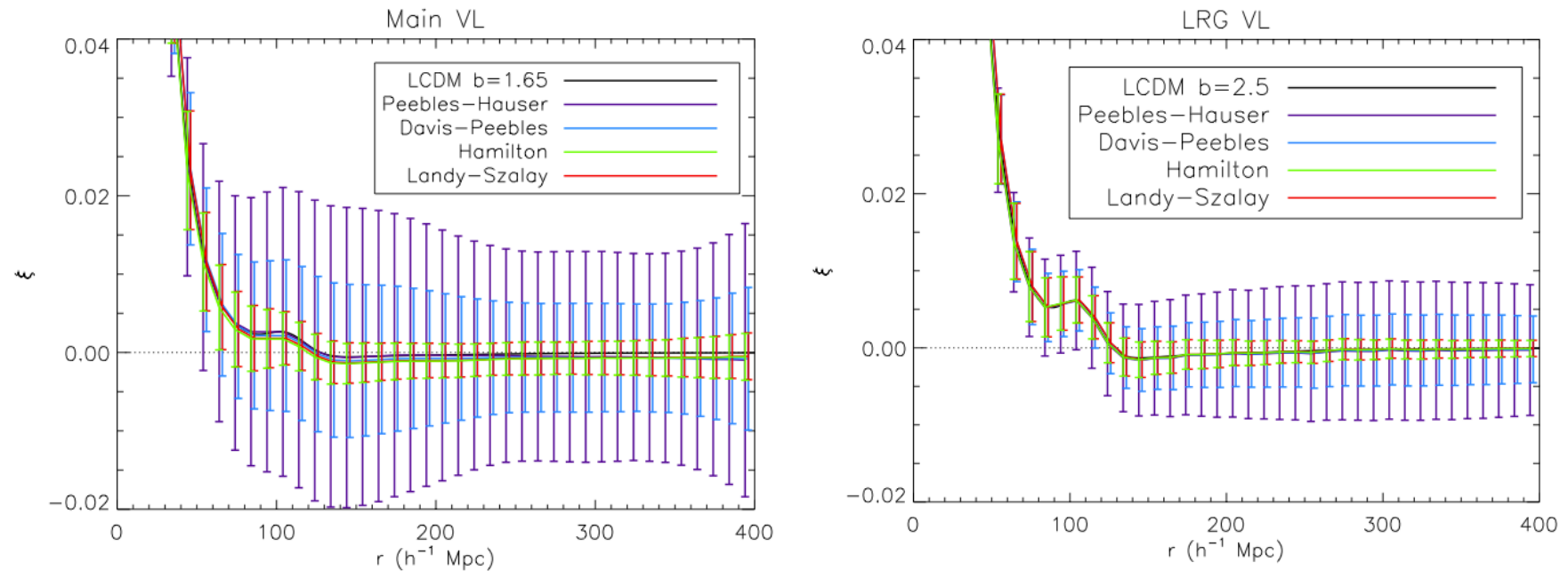


Biais non négligeable dû à la contrainte intégrale pour le Main VL



Biais négligeable pour le LRG VL

Comparaison des $\neq$ estimateurs



- Supériorité des estimateurs de Landy-Szalay et de Hamilton par rapport à Davis-Peebles et Peebles-Hauser
- Dans la suite, on se restreint à l'estimateur de Landy-Szalay

V) Modélisation statistique de ξ avec des simulations

-Nous allons valider/étudier la modélisation statistique suivante pour l'estimateur de Landy-Szalay:

$$\exists \theta \in \Theta \text{ s.t. } \hat{\xi} \sim \mathcal{N}(\xi_\theta, C_\theta)$$

Labatie et al. (2012 c)

Simulations log-normales

- Champ log-normal ρ_{LN} obtenu avec un champ Gaussien ρ_G :

$$\rho_{LN} = e^{\rho_G - \frac{\sigma_G^2}{2}}$$

- Avantages du modèle log-normal:
 - Décrit bien les probabilités de présence de galaxies $P(\delta)$ ainsi que les probabilités conditionnelles $P(\delta_2 | \delta_1)$
 - Positivité du champ assurée $\rho_{LN} > 0$ ($\neq$ du modèle Gaussien)
 - **Rapide à générer**

Modélisation de ξ pour nos simulations

- On ne garde que la dépendance en $\omega_m = \Omega_m h^2$ dans $\xi_L(r)$
- Dégradation non linéaire basée sur la formule HALOFIT + dégradation non linéaire du pic BAO:

$$P_{dw}(k) = [P_L(k) - P_{nw,L}(k)]G(k) + P_{nw,L}(k)$$

$$P_{dw,NL}(k) = \left(\frac{P_{nw,NL}(k)}{P_{nw,L}(k)} \right) P_{dw}(k)$$

- Biais des galaxies $B(r)$ étalonné à partir des simulations LasDamas
- Effet géométrique de la cosmologie fiducielle pris en compte avec paramètre de dilatation α

Simulations DR7-Full log-normales sur une grille en $\theta=(\omega_m, \alpha, b^2)$

- On utilise une modélisation avec 3 paramètres $\theta=(\omega_m, \alpha, b^2)$:

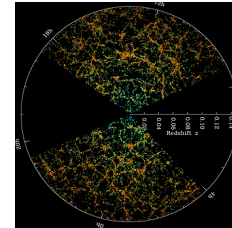
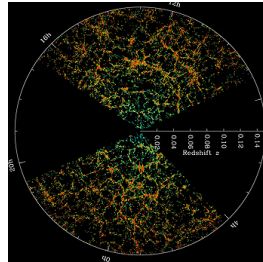
$$\xi_{\theta}(r) = b^2 B(r) \xi_{dw, \omega_m}(\alpha r)$$

- On considère une grille $\omega_m = 0.08, 0.105, 0.13, 0.155, 0.18$
- On utilise une procédure rapide pour obtenir des simulations sur une grille $0.8 < \alpha < 1.2$ ($\alpha=1$ correspond à $\Omega_m=0.25, \Omega_L=0.75$)
- On vérifie avec des simulations pour $b_1=2.5$ et $b_2=3.0$ qu'on peut modéliser analytiquement l'effet de b

Simulations DR7-Full log-normales sur une grille en $\theta=(\omega_m, \alpha, b^2)$

- Effet de α seulement géométrique:

cosmologie fiducielle
correcte



cosmologie fiducielle
incorrecte
 $\approx$ dilatation d'un facteur α

- On peut raisonner en coordonnées comobiles (et simplement varier la densité moyenne en fonction de α):

$$\bar{n}_\alpha(r) = \frac{1}{\alpha^3} \bar{n}\left(\frac{r}{\alpha}\right)$$

- Les comptages de paires en coordonnées fiducielles sont reliés aux comptages en coordonnées comobiles:

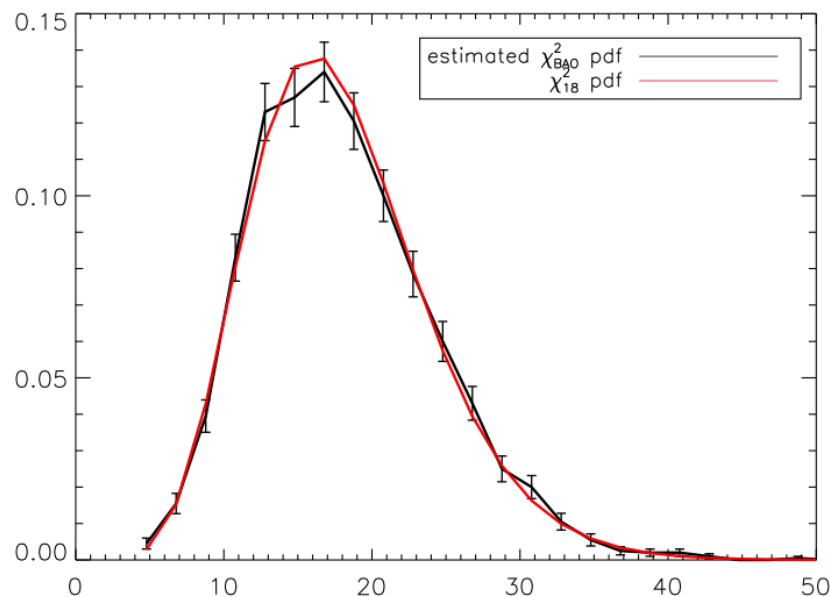
$$DD_\alpha(r) = DD_\alpha^{com}(\alpha r)$$

- **Temps total seulement $\times 4$**

Modélisation statistique de ξ :

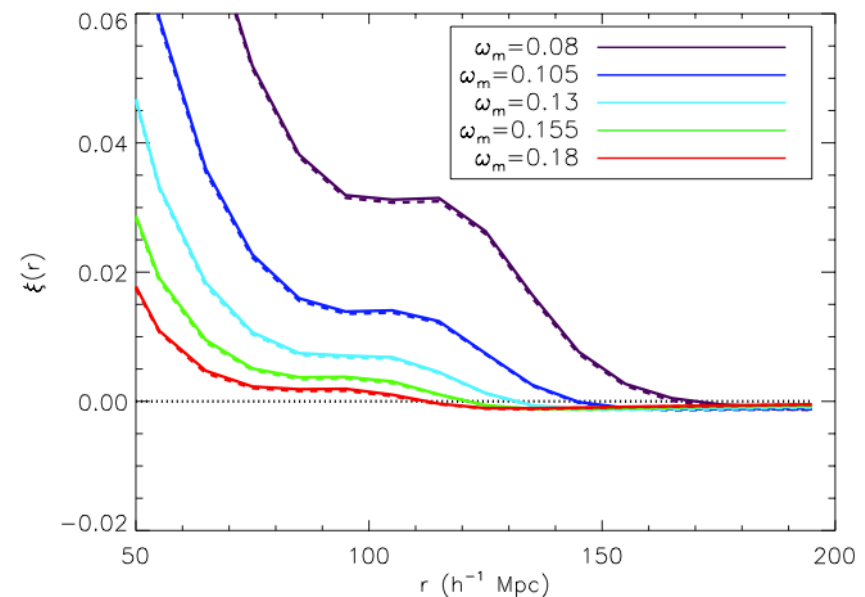
$$\exists \theta \in \Theta \text{ s.t. } \hat{\xi} \sim \mathcal{N}(\xi_\theta, C_\theta)$$

- Vérification de la Gaussianité et de l'absence de biais de l'estimateur de Landy-Szalay:



histogramme de χ^2 sur 2000 simulations (noir) et comparaison avec une loi Gaussienne (rouge):

$$\chi^2 = \langle \hat{\xi} - \xi_\theta, C^{-1}(\hat{\xi} - \xi_\theta) \rangle \rightarrow \chi^2 \sim \chi_n^2$$

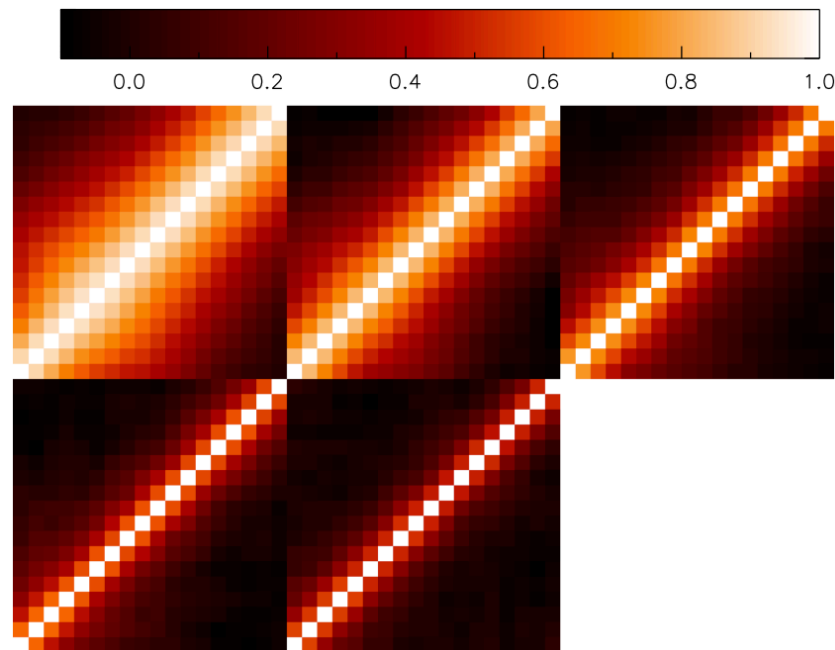


corrélation théorique (traits pleins) et moyenne de l'estimateur Landy-Szalay (tirets) sur 2000 simulations pour $\omega_m = 0.08, 0.105, 0.13, 0.155, 0.18$ (haut vers bas)

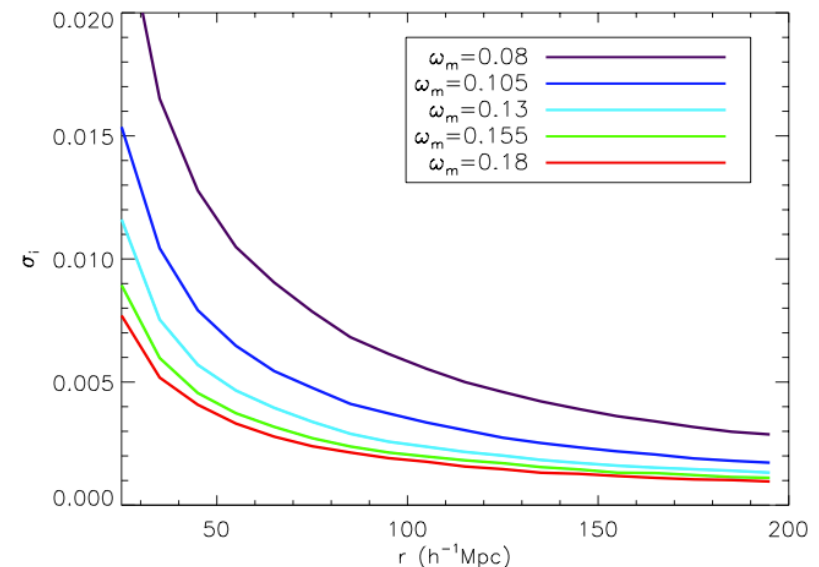
Modélisation statistique de ξ :

$$\exists \theta \in \Theta \text{ s.t. } \hat{\xi} \sim \mathcal{N}(\xi_\theta, C_\theta)$$

- Forte dépendance de C_θ par rapport à $\theta = (\omega_m, \alpha, b^2)$
- Effet de b^2 peut être bien approximé par $C_b \approx b^4 C$



matrice de corrélation ρ_θ pour $\omega_m = 0.08, 0.105, 0.13, 0.155, 0.18$ (haut gauche vers bas droite)



covariance diagonale σ_q pour $\omega_m = 0.08, 0.105, 0.13, 0.155, 0.18$ (haut vers bas)

VI) Application à la détection des BAOs

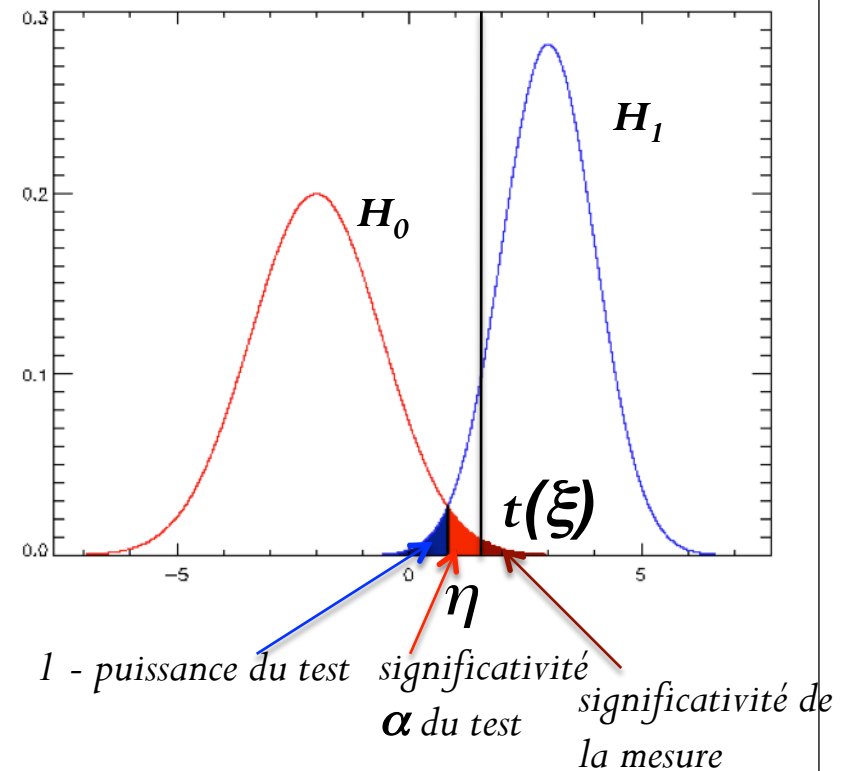
- Quelles sont les limites de la méthode classique du χ^2 pour la détection des BAOs ?
- Comment peut-on améliorer cette méthode ?

Labatie et al. (2012 b)

Tests d'hypothèses

- On veut tester 2 hypothèses H_0 and H_1 à partir d'une statistique $t(\xi)$ basée sur une mesure ξ
- Test de significativité α :
 - si $t(\xi) > \eta$ alors accepter H_1
 - si $t(\xi) \leq \eta$ alors accepter H_0
- Autre méthode standard en cosmologie:

Significativité = $P(t(\xi) > \text{valeur mesurée} \mid H_0)$



Détection des BAOs avec la méthode classique du χ^2

- Si la mesure est la fonction de corrélation :

$$\mathcal{H}_0 : \exists \theta \in \Theta \text{ s.t. } \hat{\xi} \sim \mathcal{N}(\xi_{noBAO,\theta}, C)$$

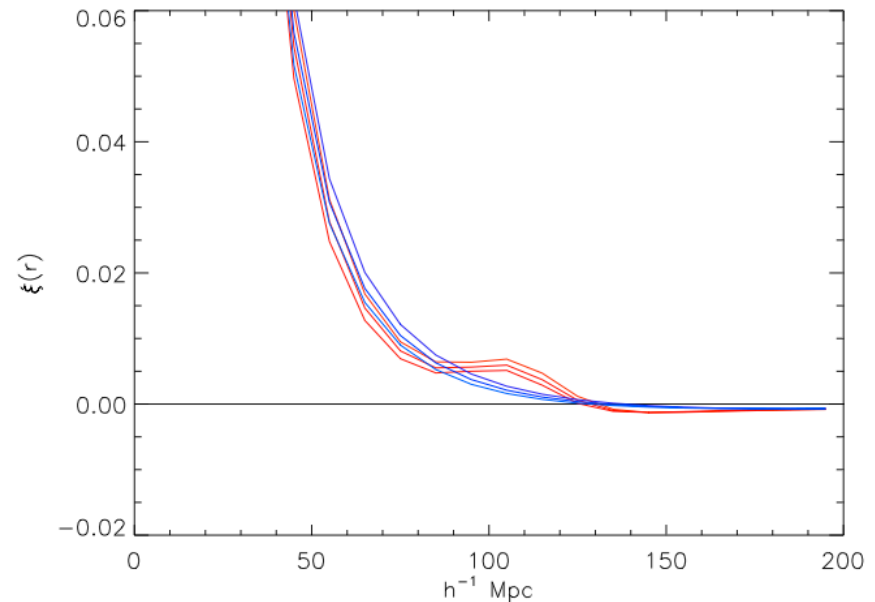
$$\mathcal{H}_1 : \exists \theta \in \Theta \text{ s.t. } \hat{\xi} \sim \mathcal{N}(\xi_{BAO,\theta}, C)$$

Suppose covariance **C** constante

- On définit la statistique du χ^2 (équivalente à la vraisemblance)

$$\chi_{BAO,\theta}^2 = \langle \hat{\xi} - \xi_{BAO,\theta}, C^{-1}(\hat{\xi} - \xi_{BAO,\theta}) \rangle$$

$$\chi_{noBAO,\theta}^2 = \langle \hat{\xi} - \xi_{noBAO,\theta}, C^{-1}(\hat{\xi} - \xi_{noBAO,\theta}) \rangle$$



différents modèles $\xi_{BAO,q}$ et $\xi_{noBAO,q}$

$$\mathcal{L}_{BAO,\theta}(\hat{\xi}) \propto \exp\left(-\frac{1}{2}\chi_{BAO,\theta}^2\right)$$

$$\mathcal{L}_{noBAO,\theta}(\hat{\xi}) \propto \exp\left(-\frac{1}{2}\chi_{noBAO,\theta}^2\right)$$

Détection des BAOs avec la méthode classique du χ^2

- $\Delta\chi^2$ = rapport de vraisemblance généralisé (optimal selon lemme de Neyman-Pearson)

$$\begin{aligned}\Delta\chi^2 &= \min_{\theta} \chi_{noBAO,\theta}^2 - \min_{\theta} \chi_{BAO,\theta}^2 \\ &= -2 \ln \left[\frac{\max_{\theta} \mathcal{L}_{noBAO,\theta}}{\max_{\theta} \mathcal{L}_{BAO,\theta}} \right]\end{aligned}$$

- La significativité peut être estimée par $\sqrt{\Delta\chi^2}.\sigma$
- **Cependant:**
 - $\sqrt{\Delta\chi^2}.\sigma$ est seulement une approximation
 - Très mauvaise approximation pour une matrice C_{θ} dépendant du modèle

Détection des BAOs avec la méthode ΔI

- Méthode force brute pour calculer une significativité correcte:

$$\mathcal{H}_0 \quad : \quad \exists \theta \in \Theta \text{ t.q. } \hat{\xi} \sim \mathcal{N}(\xi_{noBAO,\theta}, C_\theta)$$

$$P(\Delta\chi^2 \geq x \mid \mathcal{H}_0, \theta)$$

$$p(x) = \max_{\theta \in \Theta} P(\Delta\chi^2 \geq x \mid \mathcal{H}_0, \theta)$$

C_θ dépendant du modèle

- On remplace la statistique $\Delta\chi^2$ par ΔI
 ΔI = rapport de vraisemblance généralisé pour une matrice $\mathbf{C}_\theta$ dépendant du modèle

$$\Delta I = -2 \ln \left[\frac{\max_{\theta} \mathcal{L}_{noBAO,\theta}}{\max_{\theta} \mathcal{L}_{BAO,\theta}} \right] \quad \begin{array}{l} \mathcal{L}_{noBAO,\theta} \propto |C_\theta|^{-1/2} \exp \left\{ -\frac{1}{2} \left\langle \hat{\xi} - \xi_{noBAO,\theta}, C_\theta^{-1} (\hat{\xi} - \xi_{noBAO,\theta}) \right\rangle \right\} \\ \mathcal{L}_{BAO,\theta} \propto |C_\theta|^{-1/2} \exp \left\{ -\frac{1}{2} \left\langle \hat{\xi} - \xi_{BAO,\theta}, C_\theta^{-1} (\hat{\xi} - \xi_{BAO,\theta}) \right\rangle \right\} \end{array}$$

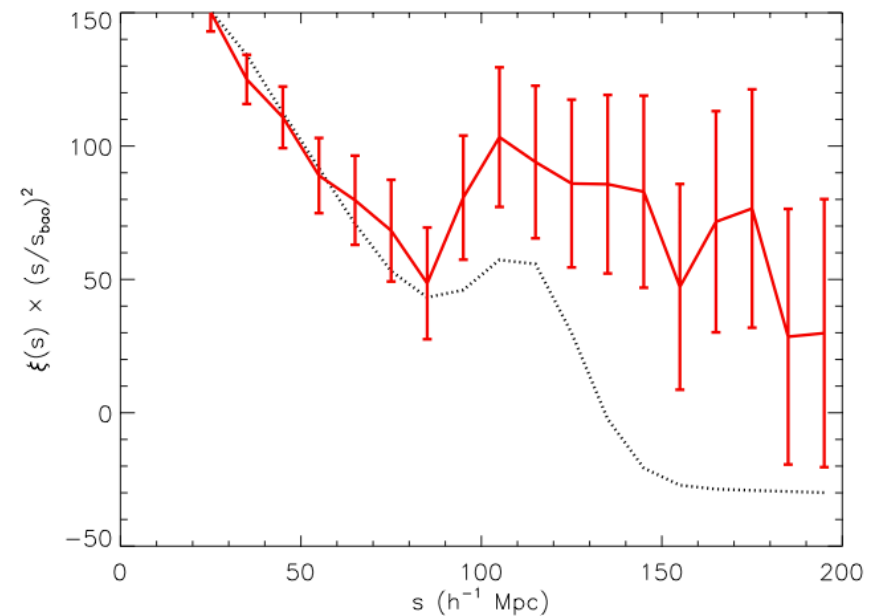
Résultats de détection des BAOs sur les données DR7-Full

- On teste les **hypothèses réalistes** suivantes:

$$\mathcal{H}_0 : \exists \theta \in \Theta \text{ t.q. } \hat{\xi} \sim \mathcal{N}(\xi_{noBAO,\theta}, C_\theta)$$

$$\mathcal{H}_1 : \exists \theta \in \Theta \text{ t.q. } \hat{\xi} \sim \mathcal{N}(\xi_{BAO,\theta}, C_\theta)$$

$\sqrt{\Delta\chi^2}$	$\sigma_{reel}(\Delta\chi^2)$	$\sigma_{reel}(\Delta l)$
1.39 σ	1.17 σ	1.43 σ



Résultats de détection des BAOs sur des simulations sous H_1

- **Significativité moyenne sous H_1 pour hypothèses avec une covariance C constante (hypothèses non réalistes):**

$$\mathcal{H}_0 \quad : \quad \exists \theta \in \Theta \text{ t.q. } \hat{\xi} \sim \mathcal{N}(\xi_{noBAO,\theta}, C)$$

$$\mathcal{H}_1 \quad : \quad \exists \theta \in \Theta \text{ t.q. } \hat{\xi} \sim \mathcal{N}(\xi_{BAO,\theta}, C)$$

$\sqrt{\Delta\chi^2}$	$\sigma_{reel}(\Delta\chi^2)$	$\sigma_{reel}(\Delta l)$
3.15σ	2.75σ	2.75σ

Résultats de détection des BAOs sur des simulations sous H_1

- **Significativité moyenne sous H_1** pour hypothèses avec une covariance C_θ dépendant du modèle (**hypothèses plus réalistes**):

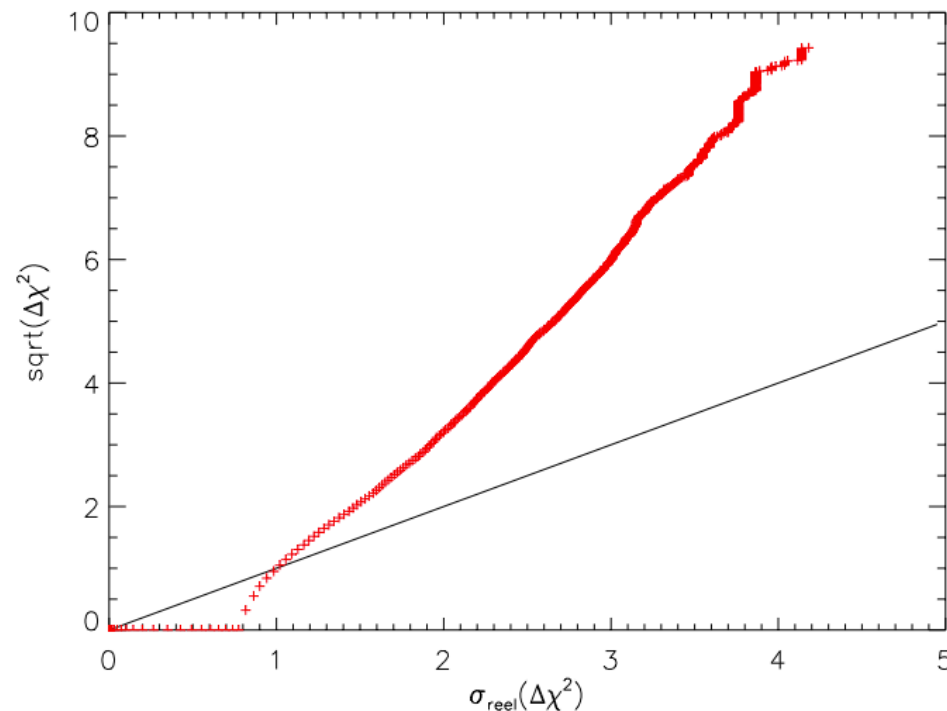
$$\mathcal{H}_0 \quad : \quad \exists \theta \in \Theta \text{ t.q. } \hat{\xi} \sim \mathcal{N}(\xi_{noBAO,\theta}, C_\theta)$$

$$\mathcal{H}_1 \quad : \quad \exists \theta \in \Theta \text{ t.q. } \hat{\xi} \sim \mathcal{N}(\xi_{BAO,\theta}, C_\theta)$$

$\sqrt{\Delta\chi^2}$	$\sigma_{reel}(\Delta\chi^2)$	$\sigma_{reel}(\Delta l)$
3.10σ	1.87σ	2.53σ

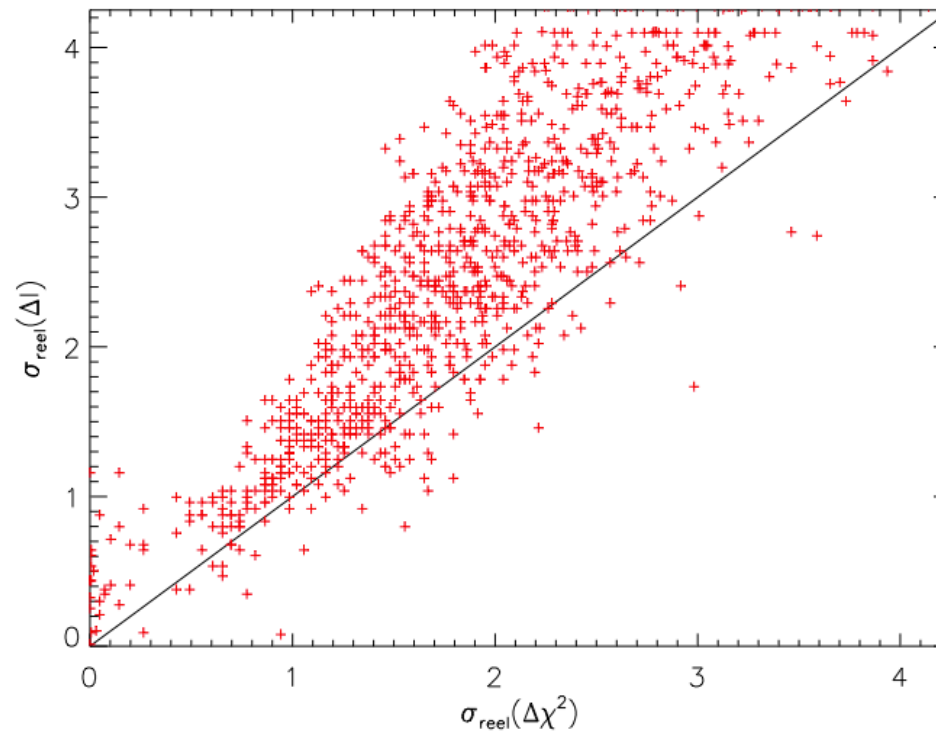
Résultats de détection des BAOs sous H_1 avec une matrice $\mathbf{C}_\theta$

- Comparaison de l'approximation $\sqrt{\Delta\chi^2}.\sigma$ avec la vraie significativité $\sigma_{\text{reel}}(\Delta\chi^2)$ pour une covariance $\mathbf{C}_\theta$ dépendant du modèle :



Résultats de détection des BAOs sous H_1 avec une matrice C_θ

- Comparaisons des significativités réelles $\sigma_{\text{reel}}(\Delta\chi^2)$ et $\sigma_{\text{reel}}(\Delta l)$ obtenues avec $\Delta\chi^2$ et Δl :



VII) Application aux contraintes cosmologiques

Quelle est l'effet de l'approximation de C_θ par une constante C pour la contrainte des paramètres cosmologiques ?

Labatie et al. (2012 c)

Contrainte des paramètres cosmologiques avec la vraisemblance

- On considère l'hypothèse suivante valide:

$$\exists \theta \in \Theta \text{ s.t. } \hat{\xi} \sim \mathcal{N}(\xi_\theta, C_\theta)$$

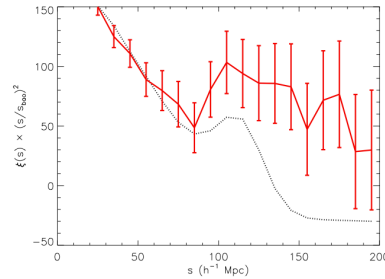
- La probabilité postérieure du paramètre θ est donnée par le théorème de Bayes:

$$p(\theta | \hat{\xi}) \propto p(\theta) \mathcal{L}_\theta(\hat{\xi})$$

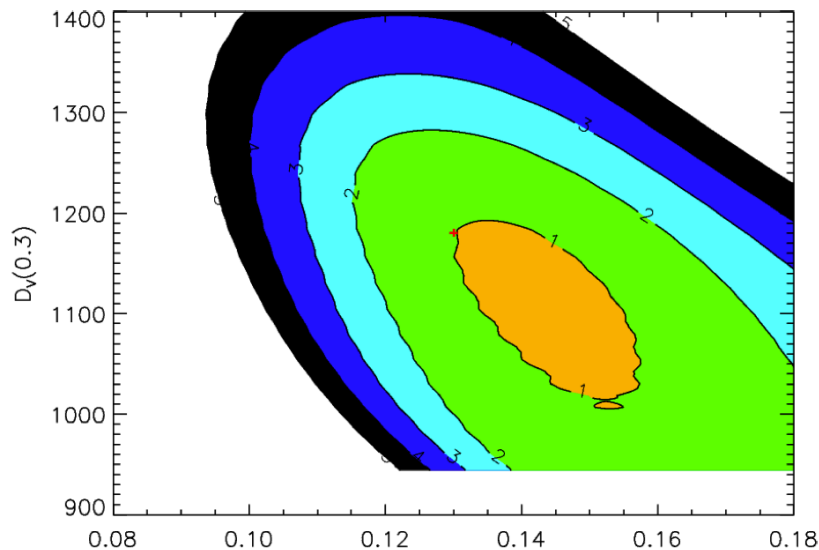
- Probabilités postérieures bi-dimensionnelles et uni-dimensionnelles obtenues par marginalisation:

$$p(\omega_m, \alpha | \hat{\xi}) = \int_{b^2} p(\omega_m, \alpha, b^2 | \hat{\xi}) db^2$$

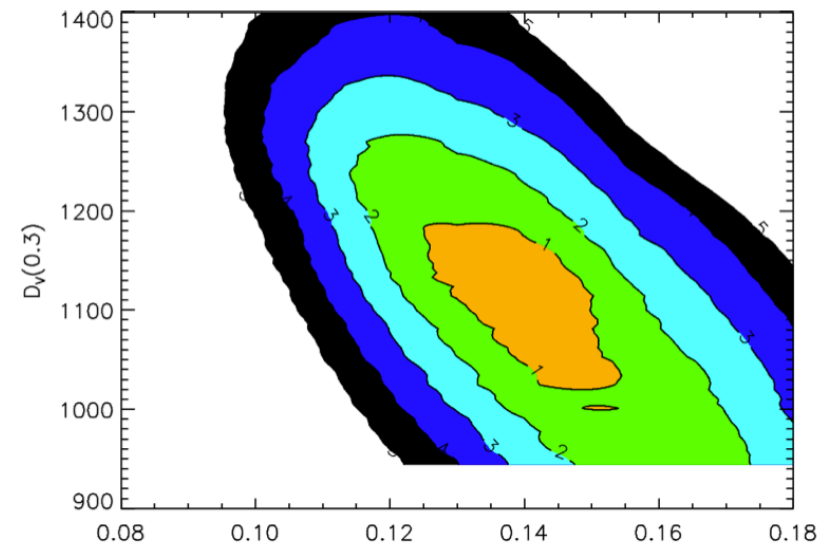
Sur le SDSS DR7-Full



$$\mathcal{L}_\theta \propto |C_\theta|^{-1/2} \exp \left\{ -\frac{1}{2} \left\langle \hat{\xi} - \xi_\theta, C_\theta^{-1} (\hat{\xi} - \xi_\theta) \right\rangle \right\}$$



Covariance ω_m constante C



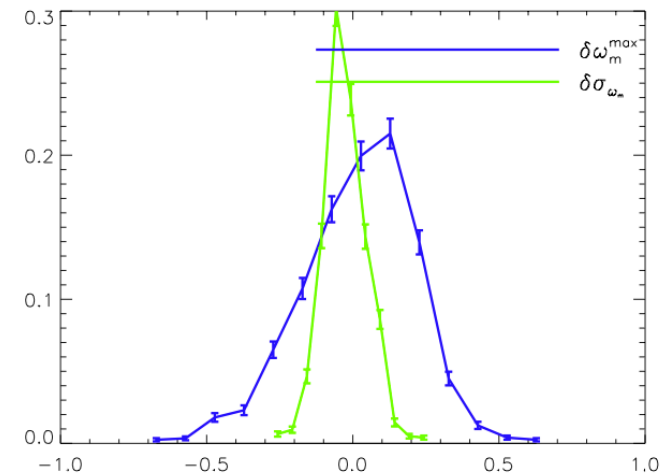
Covariance dépendant du modèle C_q

- **Léger rétrécissement** des régions de confiance pour une covariance dépendant du modèle C_θ
- Très faible **déplacement** du maximum de vraisemblance

Sur des simulations du SDSS DR7-Full

- L'approximation de C_θ par C provoque une erreur de modélisation
- **Erreurs sur le maximum de vraisemblance et sur l'intervalle à 1σ / incertitude statistique:**

$$\begin{aligned}\delta\omega_m^{\max} &= (\omega_m^{\max,C} - \omega_m^{\max,C_\theta}) / \sigma_{\omega_m}^C \\ \delta\sigma_{\omega_m} &= (\sigma_{\omega_m}^C - \sigma_{\omega_m}^{C_\theta}) / \sigma_{\omega_m}^C\end{aligned}$$



Histogramme des erreurs de modélisation /
incertitude statistique sur 2000 simulations

- **Ne peut pas être prévu a priori**
- On peut $\approx$ tenir compte de cette erreur en élargissant les intervalles à 1σ par un facteur ≈ 1.3
- **L'erreur diminue lorsque la taille des catalogues augmente**

Conclusion générale

Conclusion générale

- L'effet de la contrainte intégrale est négligeable pour les relevés de galaxies actuels
- Supériorité des estimateurs de Landy-Szalay et Hamilton
- Modélisation statistique des estimateurs:
 - Les estimateurs de Landy-Szalay sont Gaussiens et sans biais
 - Forte dépendance de la matrice $\mathbf{C}_\theta$ en fonction de θ

$$\exists \theta \in \Theta \text{ s.t. } \hat{\xi} \sim \mathcal{N}(\xi_\theta, C_\theta)$$

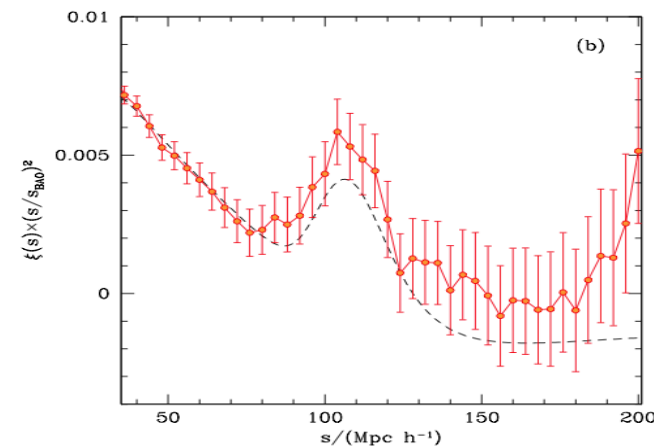
Conclusion générale

- Concernant la détection des BAOs:
 - La méthode classique du χ^2 est basée sur de fortes approximations
 - Nous avons proposé la méthode Δl comme alternative
 - Significativité correcte dans tous les cas et supériorité de Δl par rapport à $\Delta\chi^2$ pour une matrice C_θ
- L'approximation de C_θ par une constante C provoque une faible erreur sur les contraintes cosmologiques:
 - Cette erreur diminue avec la taille des relevés
 - Elle ne peut pas être prise en compte de façon exacte
 - Peut être $\approx$ prise en compte en élargissant les intervalles à 1σ par un facteur ≈ 1.3

Perspectives

- Estimation d'une matrice de covariance dépendant du modèle $\mathbf{C}_0$ pour les relevés de galaxies futurs
 - Utilisation de la méthode Δl pour détecter les BAOs
 - Réduction de l'erreur de modélisation pour les contraintes cosmologiques
- Exemple: Détection des BAOs sur BOSS DR9 CMASS

$\sqrt{\Delta\chi^2}$	$\sigma_{reel}(\Delta\chi^2)$	$\sigma_{reel}(\Delta l)$
5σ	?	?



Perspectives

- Notre méthode permet d'obtenir $\mathbf{C}_\theta$ pour 3 paramètres $\theta = (\Omega_m h^2, \alpha, b^2)$
- A l'heure actuelle, pas possible de prendre en compte tous les paramètres possibles (>6 paramètres)
 - On pourrait ne prendre en compte que les paramètres les plus importants dans $\mathbf{C}_\theta$
 - Modélisation semi-analytique de la matrice de covariance $\mathbf{C}_\theta$ qui doit encore être correctement validée ([Xu et al. \(2012\)](#))

Perspectives

- Toolbox BAOlab en C++ et IDL mise en ligne avec codes sources <http://www.cosmostat.org/baolab.html>
 - Calcul de la fonction de corrélation
 - Simulations log-normales sur une grille $\theta=(\Omega_m h^2, \alpha, b^2)$ et estimation d'une matrice C_θ
 - Détection des BAOs (méthode classique du χ^2 et méthode Δl)
 - Contraintes des paramètres cosmologiques θ

MERCI