**IV) Mocks (3 semaines : 15 mai)**

**1) Objectifs des mocks**

- pourquoi des mocks ?

--> tester l'analyse : test des systématiques (HCD, métaux…), tests de la matrice de distorsion, tests du calcul des covariances, tester les alpha, et les autres paramètres

Pas prédire la matrice de covariance

- pourquoi des GRF ?

Volume et resolution requis

brève explication des simus Nbody et Hydro, donner les pour et contre de chaque, juste pas possible

puis expliquer les GRF et donner les avantages / inconvénients

**2) Construction des mocks**

- champs de densité (QSO, Lyα)

On tire un champ gaussien dans l'espace réel

FFT pour l'avoir dans l'espace k, puis on multiplie par sqrt(P(k)/V), inverse FFT

On pourrait tirer directement dans l’espace k mais faible gain

a) les QSO :

3 transfo lognormal, à 3 z différents, avec le biais et le growth factor

A détailler

b) le lya

1 champ de densité (b=1) à z=0

- on tire les QSO

avec P \propto exp(delta)

-> Ca permet d'avoir une fonction de corrélation plus réaliste à petit r avoir CF de CAMB

Expliquer l’alternative problème qunad xi << 1 est faux, donc à petit r.

+ on déplace les QSO \propto leur vitesse (expliquer qu'on produit aussi les boites vx, vy, vz)

- densité le long des l.o.s.

on interpole a partir de chaque QSO, la densité le long de la ligne de visée

--> donner le vec lambda sur lequel c'est interpoler

+ Smoothing pour éviter l'aliasing

- densité -> transmission (FGPA) : (décrire en complexifiant au fur et a mesure)

- formule FGPA : résolution mini ~ 2 Mpc/h

- FGPA + delta\_s : expliquer le P1D\_missing (juste le principe, pas le tuning)

--> CF isotrope : on rajoute les RSD avec eta\_par (expliquer qu'on produit aussi les boites eta\_xx, ...)

Plusieurs solutions pour les RSD, mais celle ci permet d'avoir une prédiction de la CF (papier d'Andreu) avantages et inconvénients des diférentes méthodes

--> Tout ça produit les skewers, avec 0 < F < 1, dans la région Lya.

- Ajout des DLAs : poisson proportionnel à la densité. (Donner la distribution n(z) et la distrib en nhi)

On produit uniquement un catalogue de DLAs, c'est pas ajouté dans les skewers.

**3) Production des mocks**

- desi footprint : 7 chunks indépendant qui sont assemblés

- faire un schéma qui explique l'ordre de l'exécution des codes

- expliquer la pré production, puis la post-production des 100 réa,

et donner le temps CPU et la place disque de chaque étape

- mentionner aussi le BurstBuffer, les difficultés de production (le temps que ça a pris pour la pre-prod)

**4) Ajustement des paramètres**

But de l'ajustement : avoir le bon biais(z), beta(z), P1D(z) et <F>(z) pour le lya

Les QSO c'est un input direct.

Le tuning :

On choisit c(z) = beta(z) en première approx.

<F>(z) et P1D(z) donnent les bons a(z) et b(z) (besoin de plus d'info, demander à JM)

Pb : DLA dans les données, ça amplifie le P1D, donc le biais est trop grand.

--> on modifie à la main le a(z), pour avoir le bon biais(z). b(z) est modifié, mais très peu

On modifie en parallèle c(z) (pas beaucoup) pour avoir le bon beta(z).

Pour chaque (a, b, c)(z), on ajuste le P1D(delta\_s) de façon à retrouver le bon P1D à ce z.

Et on itère. A chaque itération, on calcule la prédiction en utilisant (a,b,c) et le P1D(delta\_s), et on mesure le biais et le beta avec picca (<F> est mesuré en meme temps qu'on calcule le P1D)

On fait ça pour z = 1.8, 2.2, 2.6, 3.0, 3.6

On ajuste a, b et c par des polynomes

On ajuste P1D(k, z) par un polynome

--> Montrer les paramètres ajustés

--> Montrer le biais(z), beta(z), <F>(z) et P1D(z)

--> Ceux obtenus par la prédiction

--> Ceux obtenus sur le fit de 100 raw mocks

**5) Quickquasars (Expansion des mocks)**

Expliquer rapidemment le principe (citer le papier d'alma en prep) :

reproduire les spectres de quasars observés : distribution en magnitude des QSO, 2 ou 3 spectro, la résolution des spectro, le bruit instrumental, la mesure du z des QSO (erreur du fit, dispersion des vitesse), quoi d'autre ?

on donne les skewers en input, le catalogue de QSO et de DLA

Le continuum est modélisé en utilisant les PCA

on ajoute la foret

on ajoute les DLAs, les métaux

**V) Validation des mocks (2 semaines : 29 mai)**

**1) Estimateurs de la CF, XCF**

donner les estimateurs de la CF et XCF

expliquer qu'avec les raw mocks, on a déjà F, donc on peut avoir directement delta

les cook mocks ont besoin du calcul du delta detaillé dans le chapitre 3

**2) Matrice de distorsion**

Dans le cas des cook mocks ou des données : fit du continuum --> distorsion

on corrige avec la matrice de distorsion

Donner le calcul

**3) Modélisation de la CF, XCF**

expliquer le modèle qu'on ajuste sur la CF, XCF :

d'abord le modèle ajusté sur les données

puis présenter celui qu'on ajuste sur les mocks : donner ce qu'on retire du modèle des données (les non linéarités par exemple), puis ce qu'on ajoute (gaussian smoothing)

**4) Validations des mocks**

Présenter les différents runs (eboss-0.0, eboss-0.2, eboss-0.3), et leur nombre

- Montrer le coadd ainsi que le résultat du fit pour :

- eboss-0.0 ; eboss-0.2 ; eboss-0.3

- Montrer aussi le coadd bin par bin (en z), avec ap=at=1 et comparer l'évolution en z du lya mocks/données (en fait à ce moment là j'ai besoin d'avoir parlé l'analyse en 4 bin en z de DR16 ? Ou alors je fais une référence au chapitre d'après ? C'est le même problème au moment où je parle du tuning, il faut que j'explique d'où vient l'output)

- Montrer le P1D(k), et <F>(z)

- Montrer l'auto corrélation des QSO ?

- Parler de la distrib des DLAs ?

- Expliquer que pour eboss-0.2 les DLAs sont masqués avec le vrai catalogue, et parler du DLA finder.

Montrer une CF/XCF en utilisant le catalogue de DLA produit par le finder

Comparer true catalog / finder catalog ?

**IV) Analyse des données (2 semaines : 12 juin)**

Présenter rapidement l'analyse du papier DR16.

Expliquer ce que j'ai fait moi : la même chose mais dans 4 bins en z

--> Nécessaire pour avoir biais(z) et beta(z)

Expliquer l'analyse du modèle de DLA :

- l'effet des DLA sur le lya en comparant le biais et beta dans eboss-0.0 et eboss-0.2 avec ap et at fixés

- Si j'en ai pas parlé avant : différence entre masquer les DLA à partir du true/finder catalog

- parler de l'autre modèle (pas de masque) : montrer la différence sur les mocks puis sur les données

- Donner le tableau (peut-être pas en entier) qui résume l'analyse des DLAs

Expliquer le Blinding (DESI) :

- pourquoi le blinding

- les 2 stratégies prévues

- la stratégie que j'ai implémenté et les tests qu'on a fait

Parler de l'analyse qu'a fait Jim sur l'effet de l'erreur des redshifts sur la XCF ?