

# Funciones de correlación de *shear* cósmico mediante simulaciones de *lensing*

Proyecto final

Métodos Estadísticos y Numéricos en Cosmología  
UNAM, Sem. 23-1

**Imparten:** Dra. Mariana Vargas Magaña - IF  
Dr. Sebastián Fromenteau - ICF

**Presenta:** Rafael Morales Moreno.

Diciembre 2022

## Resumen

El estudio de la llamada Estructura a Gran Escala en el universo (LSS, por sus siglas en inglés) representa actualmente un concepto que alberga interrogantes de interés en la Física contemporánea, siendo que sus orígenes estarían presumiblemente cimentados por explicaciones de la Teoría Cuántica de Campos (QFT), mientras que en el contexto cosmológico tendrían partido las explicaciones de la distribución de estructura misma y evolución de sobredensidades, en virtud de manera más directa de las observaciones. Los efectos de lentes gravitacionales resultan ser una herramienta de gran utilidad de cara a responder estas últimas preguntas.

El presente trabajo se ha enfocado al cálculo de funciones de correlación de *shear* galáctico, consecuentes del fenómeno de lentes gravitacionales débiles, o *weak lensing* (WL). Mediante datos obtenidos de simulaciones de N-cuerpos en el régimen de cielo completo, se consiguen los mapas de *lensing* de 108 realizaciones para el mismo *redshift*  $z$ , que son procesados bajo un esquema de pixelización (*healpix*), permitiendo generar las coordenadas angulares (celestes) que son correlacionadas con ayuda de la paquetería *TreeCorr*.

Consecuentemente, se muestra como resultado un conjunto de funciones de correlación  $\xi_{\pm}(\theta)$  para cada realización, calculando también la covarianza entre estos datos.

**Palabras clave:** simulaciones de N-cuerpos, *weak lensing*, función de correlación, convergencia, *shear*.

# Índice

<b>1. Introducción</b>	<b>3</b>
1.1. Motivación: ¿por qué estudiar <i>lensing</i> ( <i>shear</i> )?	3
1.2. Descripción del contenido por secciones	3
<b>2. Marco Teórico</b>	<b>4</b>
2.0.1. Formación de estructura	5
2.1. Formalismo de WL	6
2.1.1. Estimación de <i>shear</i> cósmico y modos- <i>E</i> y <i>B</i>	7
2.2. Función de correlación	8
<b>3. Descripción de datos</b>	<b>9</b>
3.1. Mapas de convergencia y <i>shear</i>	10
<b>4. Metodología</b>	<b>10</b>
4.1. Obtención de los datos de <i>lensing</i>	11
4.2. <i>Datafiles</i> (inputs): lectura de binarios	12
4.3. Conversión de datos a mapas	13
<b>5. Pipeline: jupyter nb</b>	<b>13</b>
5.1. Generación de coordenadas angulares	14
5.2. Estimador: <b>GG - TreeCorr</b>	15
5.2.1. Ejecución de los módulos y exportación de archivos ( $\xi_{\pm}$ como .txt)	16
5.3. Matriz de covarianza	17
<b>6. Resultados</b>	<b>18</b>
6.1. Relación con el curso	20

# 1. Introducción

Hacia 1919 y tras las predicciones de la Relatividad General (GR) de Einstein cuatro años antes, los primeros efectos de desviación de rayos de luz por campos de gravedad ya habían sido observados gracias al acontecimiento de un eclipse solar (Dyson et al. 1920) [1]. No fue sino hasta 1987 que efectos específicos de distorsión fuerte de imágenes por lentes extragalácticas y de gran magnitud fueron observados a causa de galaxias distantes detrás de *clusters* de mucha mayor masa (Soucaill et al.) [2]. Sin embargo, fue en 1990 (Tyson et al.) [3] que se obtuvo la primera detección de lentes gravitacionales en el régimen débil, *i.e.*, *weak lensing* (WL, a diferencia del régimen fuerte, *strong lensing*), reconociéndola en su momento como un alineamiento tangencial estadístico de dichos objetos tras su distorsión.

Entrado el nuevo milenio, recién al año 2000 (Bacon et al., Kaiser et al., Van Waerbeke et al., Wittman et al.) [4], [5], [6], [7], se mostró evidencia de WL como una distorsión coherente de galaxias medidas en “puntos ciegos”, sustentando así la existencia de *lensing* causada por la LSS del universo y, con ella misma, el acuñamiento del término **shear cósmico**.

Se entiende por *shear* galáctico (o cósmico, o simplemente *shear*) al fenómeno de distorsión de la imagen de galaxias distantes por régimen débil debido a la LSS (Kilbinger, 2015) [8]. Esta deformación se produce de manera coherente por el llamado campo de marea de las inhomogeneidades de materia a lo largo de la línea de visión (LoS por el término en inglés *line of sight*) [9].

## 1.1. Motivación: ¿por qué estudiar *lensing* (*shear*)?

A partir de medir la correlación en la forma y alineamiento de las galaxias, es posible estudiar las propiedades y la evolución de la LSS así como la geometría del universo (Kilbinger).

Las mediciones del *shear* representan una herramienta útil no solo para directa explicación de lo anteriormente mencionado, sino específicamente correlacionarlo a la naturaleza de la materia oscura y el origen del actual estado de expansión acelerada del universo.

## 1.2. Descripción del contenido por secciones

El presente escrito está designado de la manera siguiente:

La sección §2 está dedicada al estado del arte, introduciendo los conceptos y el formalismo del WL, las ecuaciones que permiten estudiarlo y las expresiones para las funciones de correlación, que son trascendentes para el entendimiento de la parte

pragmática del proyecto.

En la sección §3, una vez exploradas la teoría, terminología y definiciones, se procede a dar explicación sobre los datos utilizados en el proyecto, las funciones que cumplen y cómo fueron procesados (de manera global). Aunado a ello, se exponen los conceptos que tienen empalme con el contenido visto en el curso.

Posteriormente, en las secciones §4 y 5 se presenta la metodología utilizada, siendo que para ello se utilizaron dos códigos computacionales que cumplen dos distintas funciones en el proceso. En la sección §4 se brinda una explicación de manera más exhaustiva al procesamiento de los datos (su conversión de información binaria a mapas de *lensing* para su posterior correlación en la siguiente sección). Para ello, se ha partido de la base realizada por el trabajo de Takahashi et al. (2015) [10].

La sección §5 está dedicada a la obtención de las funciones de correlación y covarianza de los datos para una muestra de  $N$  realizaciones, a partir de la elaboración de una *notebook* propia en **Jupyter-notebook** que funciona como un *pipeline*, tomando los mapas generados por Takahashi et al. y correlacionando mediante la paquetería **TreeCorr**.

Finalmente, en la sección §6, a modo de síntesis, se muestran las gráficas resultantes de la correlación del *shear* en sus dos formas y una representación de la correlación de los datos.

**Nota:** El entregable de este proyecto consiste en **dos códigos** principales escritos en **python** y presentados en *notebooks* de **Jupyter** (`.ipynb`) y un **código adicional** demostrativo que calcula funciones de correlación de *shear* básicas (*toy model*), para entender más a fondo el funcionamiento del proceso que realizará **TreeCorr** mediante esquemas de árbol. La intención de este último es no solo clarificar el proceso (el cual, además de interesante, es muy distinto a solo relacionar posiciones de galaxias, ya que correlaciona las posiciones y deformaciones de ejes de elipses), sino también dar una mayor validez al esfuerzo puesto en desmenuzar lo que **TreeCorr** hace en un par de simples líneas).

Este proyecto no incluye inputs de datafiles anexos a las *notebooks* ya que el código mismo automatiza su descarga desde url y su posterior eliminación (además de ser archivos considerablemente pesados).

## 2. Marco Teórico

Toda la luz observada de galaxias distantes está sujeta a *lensing* gravitacional debido a las inhomogeneidades producidas por las fluctuaciones de densidad a gran escala. La distorsión que presenten en términos del tamaño y forma que adopten las galaxias determinarán la medición del *shear*, siendo que en este contexto son del orden

de apenas  $\sim 1 \times 10^{-2}$  con relación a su imagen original y que, galaxia a galaxia, el alineamiento coherente medido estadísticamente permite hacer una correlación de sus formas por pares.

Un detalle trascendente es que el *lensing* gravitacional, según Takahashi et al., no es sensible al estado dinámico de la masa en el universo en cada tipo de manifestación, sino que responde a la medición total de materia (oscura y no oscura), por lo que resulta una herramienta íntegra de los efectos gravitacionales por cualquier tipo de presencia de masa sin importar su naturaleza.

Para propagación de luz en un medio inhomogeneo, considere la métrica con perturbación a primer orden

$$ds^2 = - \left( 1 + \frac{2\Psi}{c^2} \right) c^2 dt^2 + a^2(t) \left( 1 - \frac{2\Phi}{c^2} \right) dl^2, \quad (1)$$

siendo  $c$  la velocidad de la luz propagándose en el vacío<sup>1</sup>,  $a(t)$  el factor de escala (con valor  $a(t_0) = 1$  en el presente) y  $dl^2$  la parte espacial de la métrico con coordenada comóvil  $l$  (constante con la expansión).

Los términos  $\Psi$  y  $\Phi$  en (1) son los potenciales gravitacionales, los cuales describen campos débiles, *i.e.*, satisfacen  $\Psi, \Phi \ll c^2$ . Asimismo, en ausencia de estrés anisotrópico (el cual es el caso de LSS) se cumple  $\Psi = \Phi$ .

En cuanto a elemento de línea espacial, consta de las partes radial y angulares, de la forma  $dl^2 = d\chi^2 + f_K^2(\chi)d\omega$ , donde  $\chi$  es la coordenada comóvil y  $f_K(\chi)$  está dada (según la curvatura adoptada) por

$$f_K(\chi) = \begin{cases} K^{-1/2} \sin(K^{1/2}\chi) & K > 0 \text{ (esférico)} \\ \chi & K = 0 \text{ (plano)} \\ (-K)^{-1/2} \sinh[(-K)^{1/2}\chi] & K < 0 \text{ (hiperbólico)} \end{cases} \quad (2)$$

### 2.0.1. Formación de estructura

Las fluctuaciones de densidad (dependientes del tiempo) crecieron hasta pasar de perturbaciones lineales en tiempos tempranos a no lineales. Dichas fluctuaciones son medidas a través del contraste de densidad (o sobre/sub densidad) con relación a la densidad promedio  $\bar{\rho}$  como

$$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}}, \quad (3)$$

visto más directamente como una “sobredensidad” mediante la forma  $\rho = (1 + \delta)\bar{\rho}$  (*i.e.*, qué tanto crecen, o disminuyen, en relación al promedio estadístico). Para un

---

<sup>1</sup>A pesar de la común convención  $c = 1$  tomada en la literatura habitualmente, se hace uso explícito del factor  $c^2$  par especificar el régimen débil en los potenciales y con motivo de hacer énfasis en la diferencia de orden de magnitud.

fluido perfecto (presión nula),  $\delta$  se relaciona con  $\Phi$  por la ec. de Poisson

$$\nabla^2 \Phi = 4\pi G a^2 \bar{\rho} \delta. \quad (4)$$

## 2.1. Formalismo de WL

Según GR, los fotones se propagan en geodésicas nulas dadas por un elemento de línea  $ds$  yendo a 0. De la métrica (1) se tiene entonces

$$t = \frac{1}{c} \int \left( 1 - \frac{2\Phi}{c^2} \right) dr \quad (5)$$

a lo largo del camino del rayo de luz en coordenadas propias  $r$ , en donde, según el principio de Fermat, el potencial actúa como un medio con índice de refracción variable  $n = 1 - 2\Phi/c^2$  ( $\Phi < 0$ ). De las ecuaciones de Euler-Lagrange para  $\Phi$  con  $\delta t = 0$ , se obtiene el ángulo de deflexión  $\hat{\alpha}$  entre las direcciones emitida y recibida,

$$\hat{\alpha} = -\frac{2}{c^2} \int \nabla_{\perp}^p \Phi dr. \quad (6)$$

Por otro lado, la diferencia entre el ángulo aparente  $\theta$  y el observado, permiten definir la ecuación de *lensing*

$$\beta = \theta - \alpha, \quad (7)$$

siendo  $\beta$  la posición de la fuente. Esta cantidad permite definir la **matriz (inversa) de amplificación** como el jacobiano  $\mathbf{A} = \frac{\partial \beta^2}{\partial \theta^2}$ , linealizando (7)

$$A_{ij} = \frac{\partial \beta_i}{\partial \theta_j} = \delta_{ij} - \frac{\partial \alpha_i}{\partial \theta_j}. \quad (8)$$

La deflexión puede ser escrita como el gradiente del **potencial** (bidimensional) **de lensing**  $\psi$ , tal que

$$A_{ij} = \delta_{ij} - \partial_i \partial_j \psi, \quad (9)$$

en donde

$$\psi(\theta, \chi) = \frac{2}{c^2} \int_0^\chi d\chi' \frac{f_K(\chi - \chi')}{f_K(\chi) f_K(\chi')} \Phi(f_K(\chi') \theta, \chi'). \quad (10)$$

La matriz  $\mathbf{A}$  es simétrica y se parametriza como

$$\mathbf{A} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \quad (11)$$

---

<sup>2</sup>Esta ecuación describe el mapeo lineal de la imagen con lente en coordenadas  $\theta$  a la fuente, la imagen sin lente con coordenadas  $\beta$  (Kilbinger)

en donde el término escalar  $\kappa$  se conoce como **convergencia**, mientras que  $\gamma_{1,2}$  son las componentes del objeto de espín-dos<sup>3</sup>  $\gamma = (\gamma_1, \gamma_2)$  conocido como el mencionado **shear**, siendo que representa un número complejo de la forma rectangular y polar

$$\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{2i\varphi}. \quad (12)$$

Mediante (9) es directo relacionar estas cantidades con las segundas derivadas del potencial

$$\kappa = \frac{1}{2}(\partial_1\partial_1 + \partial_2\partial_2)\psi = \frac{1}{2}\nabla^2\psi, \quad (13)$$

$$\gamma_1 = \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi, \quad (14)$$

$$\gamma_2 = \partial_1\partial_2\psi. \quad (15)$$

Por construcción, el valor esperado de ambas cantidades es nulo en promedio (*i.e.*,  $\langle\kappa\rangle = \langle\gamma\rangle = 0$ ), de modo que la primera estadística no trivial es el segundo momento. Estimadores usuales de estadística de 2° orden de WL suelen expresarse en el espacio de Fourier.

### 2.1.1. Estimación de *shear* cósmico y modos-*E* y *B*

Se le atribuye a las galaxias la cantidad intrínseca y compleja denominada elipticidad de la fuente  $\varepsilon^s$ , misma que depende y es modificada directamente como función del llamado *shear* reducido (también complejo)  $g^4$

$$g = \frac{\gamma}{1 - \kappa}, \quad (16)$$

el cual satisface las mismas propiedades y es una buena aproximación a orden lineal de  $\gamma$ .

En general, la elipticidad es  $\varepsilon = \frac{\varepsilon^s + g}{1 + g^*\varepsilon^s}$ , con  $g^*$  el complejo conjugado. En régimen de WL, la relación entre la elipticidad de la fuente y la observada está dada mediante  $\varepsilon \approx \varepsilon^s + \gamma$  (en donde también  $\langle\varepsilon^s\rangle = 0, \Rightarrow \langle\varepsilon\rangle = g$ ).

Definiendo al campo vectorial  $\mathbf{u}$  como el gradiente de la convergencia,  $\mathbf{u} = \nabla\kappa$ , y aplicando el rotacional  $\nabla \times \mathbf{u} = \partial_1 u_2 - \partial_2 u_1 = 0$ , al sustituir  $\kappa, \gamma$  en términos de  $\psi$  mediante (13). El campo de *shear* que satisface la relación anterior es llamado un campo de *modo-E* (en analogía al campo eléctrico). Respectivamente, el campo de convergencia que satisface esta condición, siendo que puede descomponerse ahora

---

<sup>3</sup>El *shear* transforma como una cantidad espín-dos: una rotación sobre un ángulo  $\pi$  es la transformación identidad de una elipse; *i.e.*, al variar  $\varphi$  entre 0 y  $2\pi$ , la elipse gira alrededor de  $\pi$  (Takahashi et al.).

<sup>4</sup>El *shear* cósmico está basado en la medición de la forma de las galaxias, de manera que la observable en cuestión no es el *shear* en sí mismo sino el *shear* reducido, dividido por un factor  $1 - \kappa$  factorizable de  $\mathbf{A}$  debido a que esta cantidad no afecta la forma de la fuente sino solo su tamaño.

en dos modos, representa al *modo-E*  $\kappa^E$  y al *modo-B*  $\kappa^B$  dados por  $\nabla^2 \kappa^E = \nabla u$  y  $\nabla^2 \kappa^B = \nabla \times u$ .

## 2.2. Función de correlación

La función de correlación de dos puntos (2PCF) de la convergencia (la cual es invariante ante traslación y rotación en analogía con el campo de densidad  $\delta$  por homogeneidad e isotropía estadísticas) es  $\langle \kappa(\vartheta) \kappa(\vartheta + \theta) \rangle$ , que en el espacio de Fourier se expresa como

$$\langle \kappa(\ell) \kappa(\ell') \rangle = (2\pi)^2 \delta_D(\ell - \ell') P_\kappa(\ell). \quad (17)$$

en términos del espectro de potencias (PS) de convergencia  $P_\kappa$  (y la distribución Delta de Dirac  $\delta_D$ ) dependiente del modo o vector de onda  $\ell$  en espacio angular (variable conjugada de  $\theta$ ).

Es importante mencionar que el PS de convergencia es igual al de *shear* ( $P_\kappa = P_\gamma$ ), y este último puede ser obtenido directamente de la observación de las elipticidades o a través de pixelización de mapas de convergencia en espacio de Fourier reconstruidas de las elipticidades mismas. Sin embargo, es más simple calcular la 2PCF del *shear*, *i.e.*, en espacio real. En términos generales, la 2PCF del *shear* puede ser estimada simplemente multiplicando las elipticidades de pares de galaxias y haciendo un promedio.

Las dos componentes de *shear* en cada galaxia son descompuestas convenientemente en la parte **tangencial**  $\gamma_t$  y **cruzada** (*cross*)  $\gamma_\times$ , definidas por su parte real e imaginaria como

$$\gamma_t = -\Re(|\gamma|e^{-2i\varphi}); \quad \gamma_\times = -\Im(|\gamma|e^{-2i\varphi}). \quad (18)$$

A partir de (18), surgen tres correlaciones de dos puntos:  $\langle \gamma_t \gamma_t \rangle$ ,  $\langle \gamma_\times \gamma_\times \rangle$  y  $\langle \gamma_t \gamma_\times \rangle$ , en donde la última se hace cero, mientras que las primeras dos permiten definir las combinaciones lineales

$$\boxed{\xi_+(\theta) = \langle \gamma_t \gamma_t \rangle(\theta) + \langle \gamma_\times \gamma_\times \rangle(\theta)}, \quad (19)$$

$$\boxed{\xi_-(\theta) = \langle \gamma_t \gamma_t \rangle(\theta) - \langle \gamma_\times \gamma_\times \rangle(\theta)}, \quad (20)$$

las **2PCF de shear**  $\xi_\pm(\theta)$ , cuyo **estimador** es

$$\hat{\xi}_\pm(\theta) = \frac{\sum_{ij} w_i w_j (\varepsilon_{t,i} \varepsilon_{t,j} \pm \varepsilon_{\times,i} \varepsilon_{\times,j})}{\sum_{ij} w_i w_j} \quad (21)$$

sobre los pares de galaxias  $(i, j)$  en las posiciones  $(\vartheta_i, \vartheta_j)$  con sus respectivas elipticidades  $\varepsilon$  (para la componente  $t$  y  $\times$ ) y pesos atribuidos  $w$ .



### 3. Descripción de datos

Takahashi et al, presentan un conjunto de datos sintéticos (simulaciones) de *lensing* gravitacional de cielo completo (*full-sky*) correspondientes a una serie de **108 realizaciones**, cada una representando el mismo experimento para distintas condiciones iniciales (o *seeds*). A su vez, una realización comprende un mapa de WL para **66** distintos valores de *redshift*, entre  $z = 0,05$  (muy cercano al tiempo presente) y  $z = 5,3$ , correspondientes a radios comóviles tomados en intervalos de  $\chi = 150$  Mpc/h entre cada mapa.

A través del sitio web creado por los autores mismos, se tiene acceso a cada conjunto de datos en formato binario para ser descargados vía `url`. Por ejemplo, para el archivo correspondiente a la primera realización (000) y al primer valor de *redshift* (0.05), se tiene:

`http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/sub1/nres12/allskymap_nres12r000.zs1.mag.dat.`

De lo anterior, es notorio que la terminación `allskymap_nres12` es fija en cada enlace de descarga, variando únicamente los siguientes caracteres, donde `r000` es la realización 0 y `zs1` es el redshift etiquetado como  $z_1$ . Asimismo, el resto de archivos tienen un enlace de descarga análogo (`r000.zs2`, `r000.zs3`, ... `r000.zs66`, `r001.zs1`, ... `r107.zs66`) (fig. 1 y 2).

realization000

weak lensing map : [zs1](#), [zs2](#), [zs3](#), [zs4](#), [zs5](#), [zs6](#), [zs7](#), [zs8](#), [zs9](#), [zs10](#), [zs11](#), [zs12](#), [zs13](#), [zs14](#), [zs15](#), [zs16](#), [zs17](#), [zs18](#), [zs19](#), [zs20](#), [zs21](#), [zs22](#), [zs23](#), [zs24](#), [zs25](#), [zs26](#), [zs27](#), [zs28](#), [zs29](#), [zs30](#), [zs31](#), [zs32](#), [zs33](#), [zs34](#), [zs35](#), [zs36](#), [zs37](#), [zs38](#), [zs66](#)

Figura 1: Ejemplo del conjunto de hipervínculos para acceder a la descarga de los data-files para cada redshift en la realización 0 y para  $N_{side} = 4096$ .

Aunado a ello, el sitio ofrece dos veces este conjunto completo de datos dependiendo del valor  $N_{side}$  que tome (parámetro de pixelización del que se dará explicación posteriormente), siendo sus posibles valores dos de ellos:  $N_{side} = 4096$  y  $N_{side} = 8192$ . A modo de resumen, el conjunto de archivos de entrada son

**$N_{side} = 4096$ :**

1. 108 realizaciones (000 - 107)

a) 66 *redshifts* (zs1 - zs66)

**$N_{side} = 8192$ :**

1. 108 realizaciones (000 - 107)

a) 66 *redshifts* (zs1 - zs66)

<u>file number</u>	<u>source redshift</u>	<u>comoving distance [Mpc/h]</u>
zs1	0.0506	150
zs2	0.1023	300
zs3	0.1553	450
zs4	0.2097	600
zs5	0.2657	750
zs6	0.3233	900
zs7	0.3827	1050
zs8	0.4442	1200
zs9	0.5078	1350
zs10	0.5739	1500
zs66	1100	9900

Figura 2: Tabla del conjunto de datos para distinto redshift y su equivalencia en distancia radial comóvil, para una misma realización y  $N_{side}$  arbitrarios.

siendo un total de 14,256 datafiles, cada uno de **3.1 GB**, en formato binario `.mag.dat`.

### 3.1. Mapas de convergencia y *shear*

Los archivos `.mag.dat` son binarios de mapas de cielo completo con información de la **convergencia**, **shear** y **rotación** de WL.

Este catálogo simulado corresponde a haber asumido también una cosmología basado en los datos de *9-year WMAP*:

$$h = 0,7, \quad \Omega_m = 0,279 \quad (= 1 - \Omega_\Lambda), \quad \Omega_{cdm} = 0,233, \quad \Omega_b = 0,046, \quad \sigma_8 = 0,82, \quad n_s = 0,97. \quad (22)$$

En general, las  $N$  realizaciones que producirán los mapas de *lensing* provienen de un mismo espectro de potencias: conjunto de partículas/objetos con  $PS_{inicial}$  en común y distintas condiciones “semilla” (configuración de las mismas) para ser evolucionadas, tal que su evolución podría resultar en una infinidad de estados finales.

## 4. Metodología

El proceso a groso modo consiste en utilizar los datos brindados por las simulaciones de Takahashi et al. para la obtención de funciones de correlación asociándole al cálculo una matriz de covarianza dadas  $N$  realizaciones (fig. 3).

Se realiza un proceso de trazado de rayos planos por múltiples lentes a través de simulaciones cosmológicas de  $N$ -cuerpos en alta resolución. En general, el proceso completo consta de dos grandes etapas (cada una con su respectivo código, fig. 4):

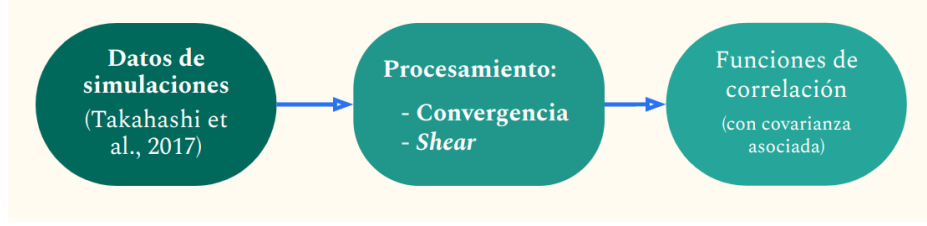


Figura 3: Proceso general del proyecto.

#### 1. WL\_maps

- a) Obtención de datos de *lensing*
- b) Lectura de archivos binarios
- c) Conversión a mapas

#### 2. Pipeline\_2PCF

- a) Generación de coordenadas angulares mediante mapas y esquema de pixelización
- b) Estimador de *shear*
- c) Matriz de covarianza

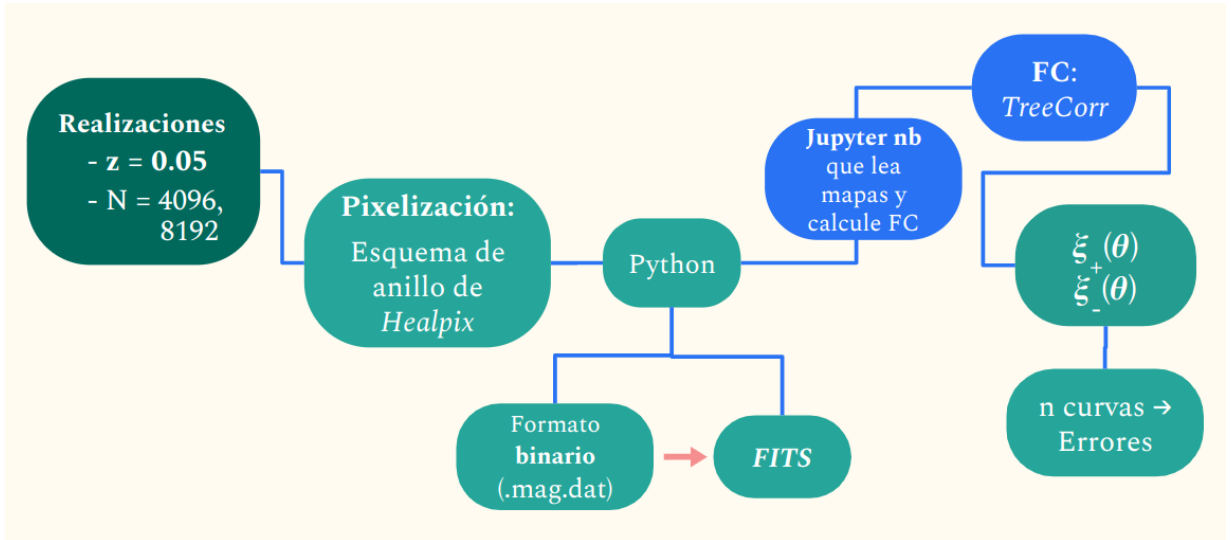


Figura 4: Diagrama de flujo detallando las fases a seguir.

### 4.1. Obtención de los datos de *lensing*

Como se ha anticipado al inicio de la sección, la fuente del artículo utilizado brinda catálogos de información de convergencia, *shear* y rotación en WL en formato binario.

Para el presente trabajo, **solamente se hace uso de la información de convergencia y shear** (que resultarán en las mismas CF) y no de rotación, **para un solo valor de  $z$  en todas sus realizaciones y tomando el parámetro  $N_{side} = 4096$ .**

Como es de esperarse, debería existir una forma más eficiente de acceder a los 108 archivos de datos, siendo que se escribió un pequeño *script* que fue anexado al código original para la descarga automática de los mismos (y su misma eliminación al terminar para evitar duplicación cada vez que se pruebe el código) (fig. 5).

### Script to automatically download data files (inputs) from the url of Takahashi et al.

[http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky\\_raytracing/nres12.html](http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/nres12.html)

```
import urllib.request
import time

# defining the url string as the first part of the link
string1 = 'http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/sub1/nres12/allsky'

start = time.time()
for i in range(1): # 108 realizations in total (JUST ONE z each)
    if i < 10:
        datafiles_url = string1 + '00' + str(i) + '.zsl.mag.dat'
        local_file = 'allskymap_nres12r00' + str(i) + '.zsl.mag.dat' # downloads and save in the
        urllib.request.urlretrieve(datafiles_url, local_file)
    if 10 <= i <= 99:
        datafiles_url = string1 + '0' + str(i) + '.zslmag.dat'
        local_file = 'allskymap_nres12r0' + str(i) + '.zsl.mag.dat'
        urllib.request.urlretrieve(datafiles_url, local_file)
    if 99 < i:
        datafiles_url = string1 + str(i) + '.zslmag.dat'
        local_file = 'allskymap_nres12r' + str(i) + '.zsl.mag.dat'
        urllib.request.urlretrieve(datafiles_url, local_file)

end = time.time()
```

Figura 5: *Script* para la automatización de descarga (y eliminación posterior a su uso) de cada datafile requerido.

## 4.2. *Datafiles* (inputs): lectura de binarios

Una vez descargados los archivos y generadas la variable en *python* que se le asocia a cada uno, se procede a tomar lectura de la información binaria (fig. 6). Cabe mencionar que el código proporcionado originalmente es un archivo de texto que funciona únicamente para un solo input (*i.e.*, una sola realización en un solo *redshift*), por lo que fue necesaria la implementación de una secuencia y la creación de un módulo que tomara absolutamente todas las realizaciones una a una (fig. 7).

Dentro del módulo, existe una secuencia *with* de lectura de archivo a archivo, encargada de definir las variables de pixelización *nside*, *npix*, *rec* para cada uno de los archivos, necesarias para la creación de mapas en el esquema de *healpix*.

## Reading binary data

```
import numpy as np
import healpy as hp

# input file
filename = 'allskymap_nres12r001.zs16.mag.dat' # just ONE file alone (read one by one)

# define array (why those elements?)
skip = [0, 536870908, 1073741818, 1610612728, 2147483638, 2684354547, 3221225457]

load_blocks = [skip[i+1]-skip[i] for i in range(0, 6)] # array of shape (6, ) of blocks for
```

Figura 6

## Constructing an array from data in *binary* file

From `xxxx.zs##.mag.dat` file takes information in order to define the next arrays:

### Reading binary data

```
def make_experiment(filename, num_realization):
    # define array
    skip = [0, 536870908, 1073741818, 1610612728, 2147483638, 2684354547, 3221225457]
    load_blocks = [skip[i+1]-skip[i] for i in range(0, 6)] # array of shape (6, ) of blocks

    with open(filename, 'rb') as f: # allows simplifying writing
        rec = np.fromfile(f, dtype='uint32', count=1)[0] # define array of specific variable
        nside = np.fromfile(f, dtype='int32', count=1)[0] # (but now in particular just take
        npix = np.fromfile(f, dtype='int64', count=1)[0] # and takes the index [0], so it's
        rec = np.fromfile(f, dtype='uint32', count=1)[0]
        print("nside:{} npix:{}".format(nside, npix)) # prints the corresponding N_side

        rec = np.fromfile(f, dtype='uint32', count=1)[0] # why again?

    ##### CONVERGENCE part #####
```

Figura 7: Módulo para la lectura de cada uno de los datafile (108 realizaciones).

## 4.3. Conversión de datos a mapas

Dentro de la misma secuencia de construcción (fig. 7), se tiene un loop de llenado para la información respectiva de convergencia (`kappa`) y ambas componentes del shear (`gamma1`, `gamma2`) (fig. 8).

Finalmente, esta información (`kappa`, `gamma1`, `gamma2`) es escrita en un formato de mapa a través de `healpix` en formato `.fits` (fig. 9), misma que es la información de salida del módulo en la subsección anterior.

## 5. Pipeline: jupyter nb

Una vez obtenidos los mapas del cielo, se procede a tomar esta información para trabajar la correlación. Por motivos de poder de cómputo se optó por hacer el proceso consecuente (que es el más pesado) en un cuaderno distinto, el cual cumple las

```

##### CONVERGENCE part #####

kappa = np.array([])    #empty array
r = npix                # redefine the previous variable from f

for i, l in enumerate(load_blocks): # similar to for+range but with an iterator (two ar
    blocks = min(l, r) # define the min. between the values of load_blocks
    load = np.fromfile(f, dtype='float32', count=blocks) # now takes n=blocks count (# o
    np.fromfile(f, dtype='uint32', count=2) # takes 2 items from f
    kappa = np.append(kappa, load) # add these values to the initial empty array
    r = r-blocks # takes away the minimum

    if r == 0: # we avoid the case r=0
        break
    elif r > 0 and i == len(load_blocks)-1:
        load = np.fromfile(f, dtype='float32', count=r) # def. array of 'r' elements
        np.fromfile(f, dtype='uint32', count=2)
        kappa = np.append(kappa, load)

```

Figura 8

```

print('loading completed')

name_kappa = 'output_k_' + str(num_realization) + '.fits'
name_gamma1 = 'output_g1_' + str(num_realization) + '.fits'
name_gamma2 = 'output_g2_' + str(num_realization) + '.fits'

hp.fitsfunc.write_map(name_kappa, kappa)
hp.fitsfunc.write_map(name_gamma1, gamma1)
hp.fitsfunc.write_map(name_gamma2, gamma2)

return kappa, gamma1, gamma2, npix

```

Figura 9: Escritura de los mapas de convergencia y componentes del *shear*.

funciones de subir y leer los archivos **.fits** uno a uno (mediante un módulo nombrado *function1*, fig. 10), realizar un esquema de pixelización mediante información de resolución angular, generar las coordenadas celestes correspondientes a la información del *shear1,2*, y realizar un catálogo de **TreeCorr** que asocie las posiciones (ángulos dados). Finalmente grafica un ejemplo de realización de las funciones  $\xi_+(\theta)$  y  $\xi_-(\theta)$  y exporta esta información en un archivo de formato **.txt** para su futura correlación y covarianza.

## 5.1. Generación de coordenadas angulares

El módulo **healpix** disponible para Python es utilizado para realizar esquemas de pixelización en tratamientos de espacios curvos, particularmente en las coordenadas

## Reading *.fits* maps

Read a healpix map from a fits file. Partial-sky files, if properly identified, are expanded to full size and filled with UNSEEN

```
def function1(num_realization):
    gamma1=hp.fitsfunc.read_map(filename = 'output_g1_' + str(num_realization) + '.fits')
    gamma2=hp.fitsfunc.read_map(filename = 'output_g2_' + str(num_realization) + '.fits')
```

Figura 10: Lectura de mapas de lensing (de la forma `gamma1.fits`, `gamma2.fits`) para N realizaciones solicitadas por el usuario.

esféricas  $\theta, \phi$ . Como parte de esta *notebook*, fue necesaria la creación de un segundo módulo o función *function2* (fig. 11) que utilizara dicho esquema, definiendo para su funcionamiento las variables de pixelización  $N_{side}$  (resolución angular de la forma  $2^n$  a partir del número de pixeles deseado).

Habiendo definido este esquema, se procede a generar las coordenadas para esta pixelización, definiendo las coordenadas celestes y dotándolas de la información en `gamma1,2` obtenida previamente (fig. 12).

```
NSIDE = 4096 # angular resolution size

tth, phi = hp.pix2ang(nside = NSIDE, ipix = range(hp.nside2npix(NSIDE)))# in radians
# << hp.nside2npix >> gives the number of pixels for the given nside.

if num_realization == 1:
    print(len(tth)) # number of values/elements in theta array

# Defining limit values of the angle in arcmin and its conversion to rad
minArcmin = 5.
maxArcmin = 200.

minRad = minArcmin*np.pi/60/180 # converting values to radian
maxRad = maxArcmin*np.pi/60/180

if num_realization == 1:
    print("Chosen angular limit values as: \n \n minRad={:.5}, maxRad={:.5}".format(minRad,maxRad))

nInputSize = len(tth) # defining size
nSample = int(nInputSize/2) # max is len(tth) # two #theta for one sample

if num_realization == 1:
    print("Taking", nSample, "points out of", nInputSize )
```

Figura 11: Lectura de mapas de lensing (de la forma `gamma1.fits`, `gamma2.fits`) para N realizaciones solicitadas por el usuario.

## 5.2. Estimador: GG - TreeCorr

La paquetería `TreeCorr` alberga dentro de sus funciones directamente la herramienta para calcular correlaciones de *shear* (`GGCorrelation`, *i.e.*, Gamma - Gamma).

```

nDec, nRac, nG1, nG2 = [], [], [], [] # empty arrays to be filled w/ the Declin., right ascension and shear1,2

#rng = np.random.default_rng(2021)
#nii = rng.choice(len(tht), size=nSample, replace=False)

np.random.seed(2022) # taking initial conditions in order to make random realizations 'reproducible'

# Choosin randomly unrepeated terms between (0 - nInputSize)
nii = np.random.choice(nInputSize, nSample, replace=False) # Which has the syntaxis: random.choice(a, size=None, 1
# where a random sample is generated as np.arange(0,a)

for ii in nii:
    nDec.append(tht[ii]) # declination
    nRac.append(phi[ii]) # right ascension
    nG1.append(gamma1[ii]) # shear 1
    nG2.append(gamma2[ii]) # shear 2

del tht, phi, gamma1, gamma2 # ereasing data (to reset)
return nDec, nRac, nG1, nG2, nSample

```

Figura 12: Generación de coordenadas celestes con Healpix.

Para ello, es necesario definir un catálogo que tome por entrada las variables de coordenadas esféricas (en este caso las celestes: ascensión recta y declinación), así como las componentes del *shear* mismo.

```

def function2(nDec,nRac,nG1,nG2,nSample,num_realization):
    minArcmin = 5.
    maxArcmin = 200.

    minRad = minArcmin*np.pi/60/180 # converting values to radian
    maxRad = maxArcmin*np.pi/60/180

    # Now using T r e e C o r r package to compute CF

    # building catalogue passing arrays for each piece of information
    cat = treecorr.Catalog(\
        ra = nRac, dec=nDec, g1=nG1, g2=nG2, ra_units="rad", dec_units="rad") # catalogue defined

    #cat = treecorr.Catalog(\
    #    ra=phi,dec=tht,g1=gamma1,g2=gamma2,ra_units="rad",dec_units="rad") # same with spherical coord. instead of celestial

    # Now creating the GG correlation (shear - shear corr.)
    GG = treecorr.GGCorrelation(min_sep=minRad, max_sep=maxRad, verbose=2, nbins=20)
    # where: min_sep - minimum separation being considered
    #           max_sep - analogous

```

Figura 13: Uso de la paquetería TreeCorr para la creación del catálogo GG (gamma-gamma, *i.e.*, *shear*).

### 5.2.1. Ejecución de los módulos y exportación de archivos ( $\xi_{\pm}$ como .txt)

Finalmente en cuanto a lo que este proceso de obtención consierne, un último módulo es creado para guardar las CF en arreglos que contengan tanto la coordenada independiente  $\theta$  como las componentes + y - del *shear* (fig. 14). De esta manera, se ejecutan las funciones y se exportan los archivos de texto (`xi_pm_i.txt`) (fig. 15) que serán utilizados para el cálculo de la covarianza.



```

def save_output(outputfilename, GG):
    salida = np.zeros((len(GG.xim),4)) # array of zeros to be filled w/
    #print(salida) # Prints
    thtArr=np.array(GG.meanr)*60*180/np.pi
    for ii in range(len(GG.xim)):
        salida[ii,0] = thtArr[ii]
        salida[ii,1] = GG.xip[ii] # 'plus' component
        salida[ii,2] = GG.xim[ii] # 'minus' component
        salida[ii,3] = GG.npairs[ii]

    #print(salida)

    #outputfilename='xi_pm.txt'

    header = "# 1.meanr  2.xi_plus  3.xi_minus  4.npairs"

    with open(outputfilename, 'a') as f:
        f.write(header + "\n")
        np.savetxt(f, salida)
        f.close

```

Figura 14: Módulo destinado al guardado de archivos .txt.

```

# Running module 'create_output' with every single file

start_time = time.time() # Computing time

for i in range(1,109): # from i=1 until i=109 for the 108 realizations in total (JUST ONE z each)
    if i < 10:
        outputfilename = 'xi_pm_'+str(i)+".txt"
        nDec, nRac, nG1,nG2, nS = function1(i)
        GG = function2(nDec,nRac,nG1,nG2,nS,i)
        #create_output(i, outputfilename)
        save_output(outputfilename, GG)
    if 10 <= i <= 99:
        outputfilename = 'xi_pm_'+str(i)+".txt"
        nDec, nRac, nG1,nG2, nS = function1(i)
        GG = function2(nDec,nRac,nG1,nG2,nS,i)
        # create_output(i, outputfilename)
        save_output(outputfilename, GG)

end_time = time.time()
print('Time:', end_time - start_time, 's.')
print('output file will be xi_pm_i.txt,\n where i is the realization number from 0 to 107')

NSIDE = 4096
ORDERING = RING in fits file
INDXSCHM = IMPLICIT

```

Figura 15: Ejecución de las funciones (módulos) y obtención salidas ( $\xi_{\pm}$ ).

### 5.3. Matriz de covarianza

Se realizó también un cálculo de las matrices de covarianza y correlación asociadas a los datos, en específico, a la información referente a las funciones de correlación del *shear*  $\xi_{\pm}(\theta)$ . Entendiendo a la **matriz de covarianza** como la generalización de la covarianza en  $n$  dimensiones ( $N$  conjuntos de datos), tal que resulte un arreglo de  $n \times n$  de la forma

$$\Sigma = \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \cdots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \cdots & \text{Var}(X_n) \end{bmatrix}$$

en donde claramente cada elemento de la misma representa a la covarianza definida por

$$\text{Cov}_{i,j} = \text{Cov}(x_i, x_j) = \frac{1}{N} \sum_{k=1}^n (x_k - \bar{x}_{i(k)}) (x_k - \bar{x}_{j(k)}) .$$

Cabe notar que la autocovarianza es la varianza en sí misma, es decir,  $\text{Cov}_{i,i} = \text{Cov}(x_i, x_i) = \text{Var}(x_i) \equiv \sigma_{x_i}^2$ . Análogamente y solo difiriendo en los términos de normalización (denominador), la **matriz de correlación** es la matriz con elementos definidos por la correlación

$$r_k = \frac{\sum_{i=1}^{N-k} (x_i - \bar{x}) \cdot (y_{i+k} - \bar{y})}{\sqrt{\sum_{i=1}^{N-k} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=k+1}^N (y_i - \bar{y})^2}}$$

## 6. Resultados

Como productos del presente proyecto se han obtenido las gráficas correspondientes a las funciones  $\xi_{\pm}(\theta)$  para 15 realizaciones, mostrando aquí la primera de ellas y la curva promedio (figs. 16 y ??). Asimismo, por motivos de poder de cómputo y almacenamiento, únicamente una pequeña muestra de las realizaciones pudo ser tomada: 15 funciones de correlación “+” y 15 “-”, para una matriz de covarianza y otra de correlación de 15×15 para cada componente.

Por razones lógicas, el resultado de la correlación muestra ruido considerable, especialmente en los primeros elementos; sin embargo, la correlación sí es visualmente notable al verse un rastro característico en su diagonal (fig. 18). Un resultado análogo fue obtenido para la covarianza (fig. 19).

**Nota:** cabe resaltar que el código está diseñado para correr el total de realizaciones, i.e., funciona para N=108 o cualquiera ingresado por el usuario en virtud del poder de cómputo del que disponga y el almacenamiento para los datos descargados.

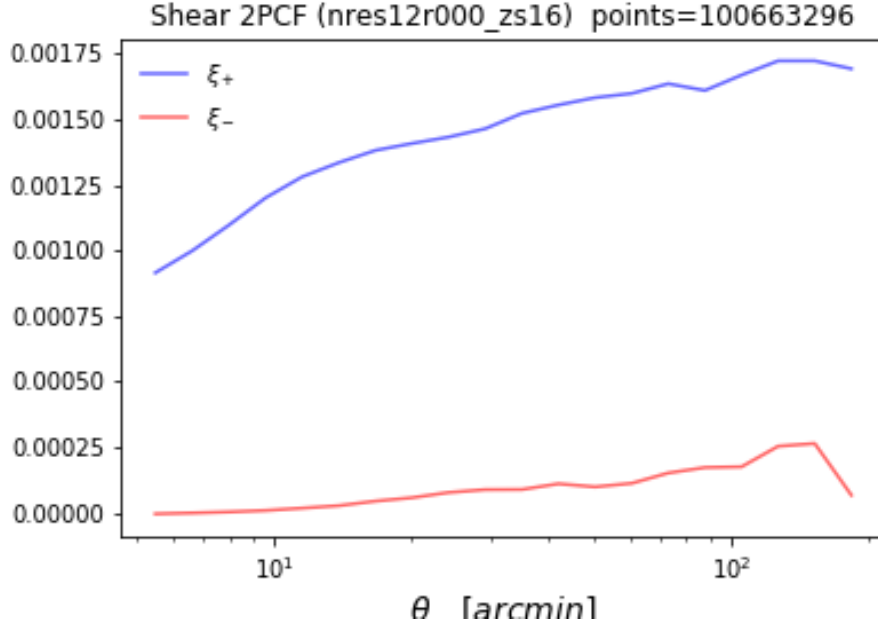


Figura 16: Función de correlación del *shear*  $\xi_{\pm}$  para la realización *zs000* en el redshift  $z=0.05$ .

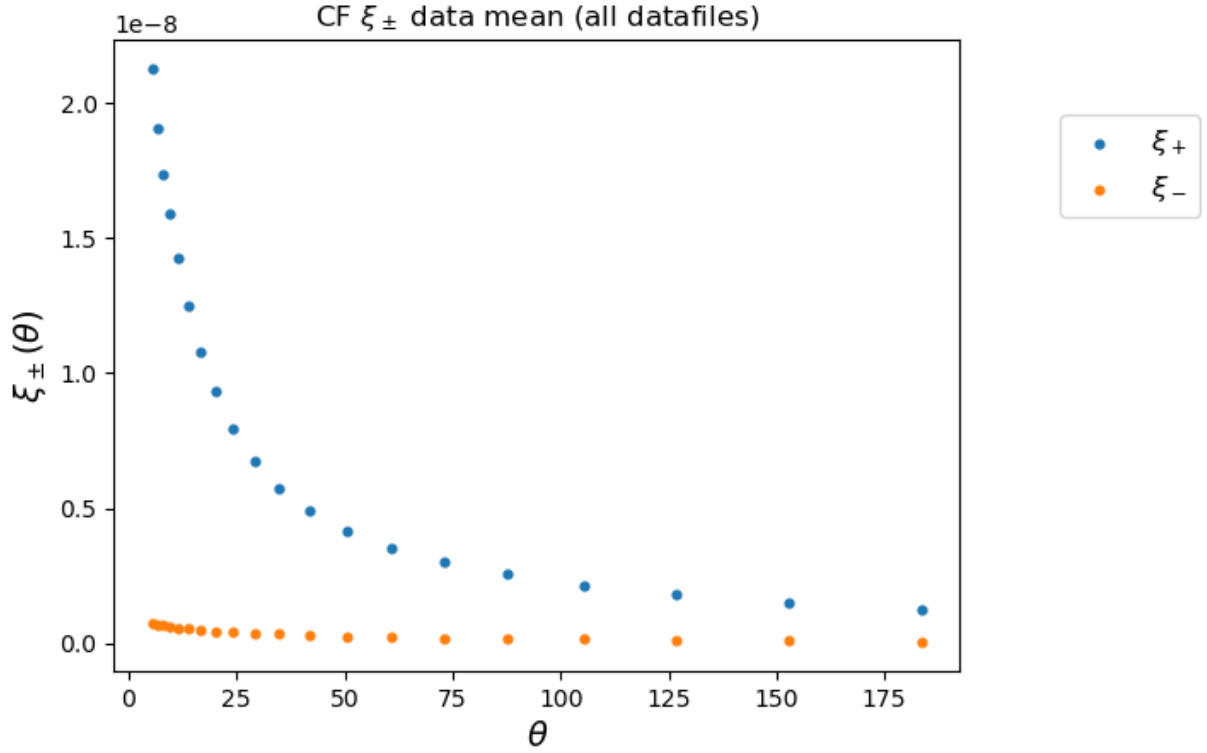


Figura 17: Promedio de todas las realizaciones, *i.e.*,  $\xi_{\pm}^{mean}$  desde 000 hasta 014.

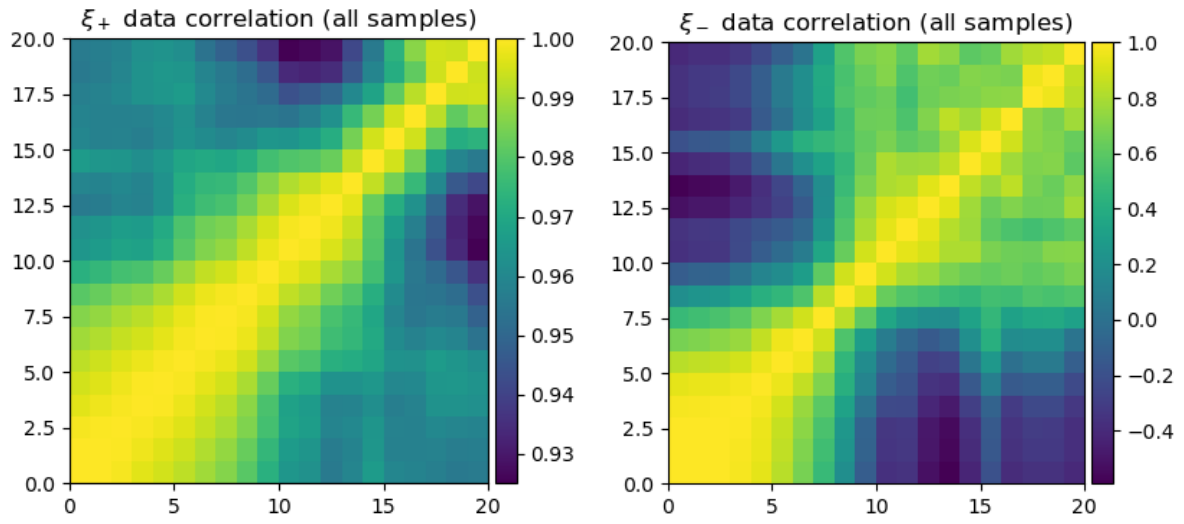
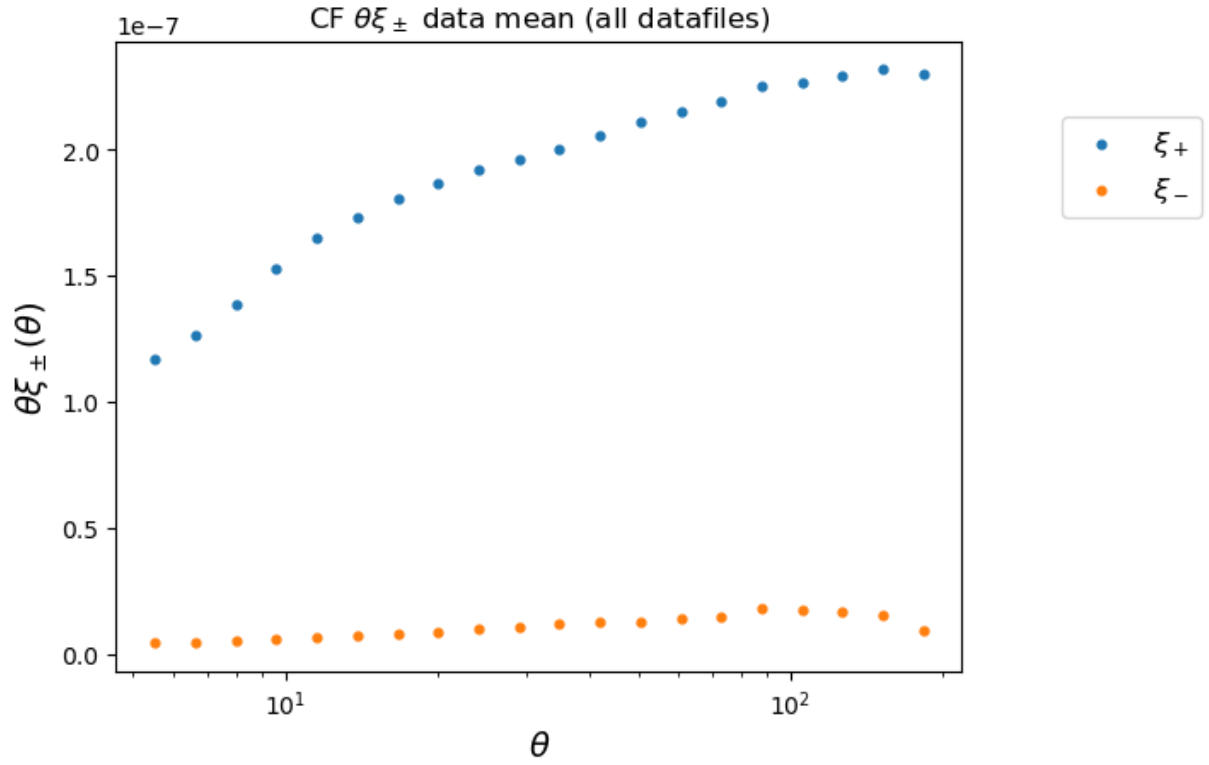


Figura 18: Matriz de correlación de las 15 realizaciones de  $\xi_+$  y  $\xi_-$ .

## 6.1. Relación con el curso

1. *N-point statistics*
2. Funciones de correlación de 2-pts

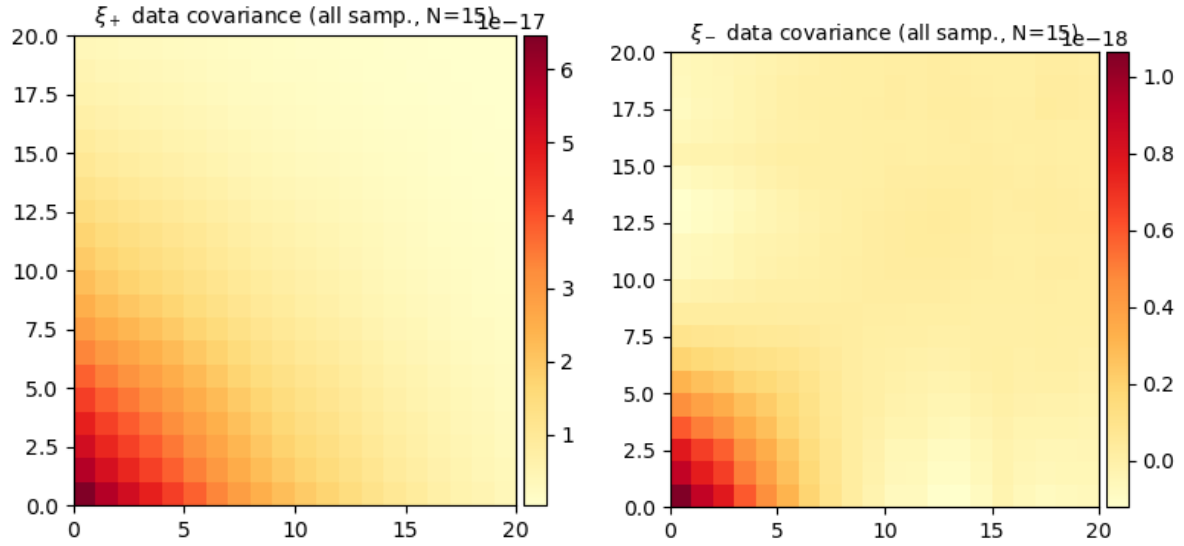


Figura 19: Matriz de covarianza para los mismos datos y casos.

3. Consecuencias de las distorsiones de distribución de galaxias (espacio real en WL  $\Rightarrow$  espacio  $z$  en analogía con RSD)
4. Estimación de errores: matriz de covarianza

## Referencias

- [1] F. W. Dyson, A. S. Eddington, and C. Davidson, “Ix. a determination of the deflection of light by the sun’s gravitational field, from observations made at the total eclipse of may 29, 1919,” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 220, no. 571-581, pp. 291–333, 1920.
- [2] G. Soucail, B. Fort, Y. Mellier, and J. Picat, “A blue ring-like structure in the center of the a 370 cluster of galaxies,” *Astronomy and astrophysics*, vol. 172, pp. L14–L16, 1987.
- [3] J. A. Tyson, F. Valdes, and R. Wenk, “Detection of systematic gravitational lens galaxy image alignments-mapping dark matter in galaxy clusters,” *The Astrophysical Journal*, vol. 349, pp. L1–L4, 1990.
- [4] D. J. Bacon, A. R. Refregier, and R. S. Ellis, “Detection of weak gravitational lensing by large-scale structure,” *Monthly Notices of the Royal Astronomical Society*, vol. 318, no. 2, pp. 625–640, 2000.
- [5] N. Kaiser, G. Wilson, and G. A. Luppino, “Large-scale cosmic shear measurements,” *arXiv preprint astro-ph/0003338*, 2000.

- [6] L. Van Waerbeke, Y. Mellier, T. Erben, J. Cuillandre, F. Bernardeau, R. Maoli, E. Bertin, H. Mc Cracken, O. L. Fevre, B. Fort, *et al.*, “Detection of correlated galaxy ellipticities on cfht data: first evidence for gravitational lensing by large-scale structures,” *Arxiv preprint astro-ph/0002500*, 2000.
- [7] D. M. Wittman, J. A. Tyson, D. Kirkman, I. Dell’Antonio, and G. Bernstein, “Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales,” *nature*, vol. 405, no. 6783, pp. 143–148, 2000.
- [8] M. Kilbinger, “Cosmology with cosmic shear observations: a review,” *Reports on Progress in Physics*, vol. 78, no. 8, p. 086901, 2015.
- [9] S. Dodelson and F. Schmidt, *Modern cosmology*. Academic Press, 2020.
- [10] R. Takahashi, T. Hamana, M. Shirasaki, T. Namikawa, T. Nishimichi, K. Osato, and K. Shiroyama, “Full-sky gravitational lensing simulation for large-area galaxy surveys and cosmic microwave background experiments,” *The Astrophysical Journal*, vol. 850, no. 1, p. 24, 2017.
- [11] A. Avilés. Cosmological weak lensing of galaxy sources. Curso de Primavera 2022 en “Weak Lensing: theory & estimators”, IAC (notas), 2022.