Note for Correlation Functions

Sehighs

Aug. 2025

1 Basic of 2PCF

The 2-point correlation function as used in astrophysics describes one way in which the actual distribution of galaxies deviates from a simple Poisson distribution. There are other descriptors like three point correlation functions, the topological genus and so on.

For Galaxy field, we treat it as a discrete point field.

First, in random point field: Given a point at r_0, and the probability of found another point in $\vec{r}$ is just:

$$dP(r) = \bar{n}dV = \bar{n}\,4\pi r^2 dr \qquad \bar{n} \equiv \langle n \rangle, \tag{1}$$

the $\langle\rangle$ is the mean of the space, so the $\bar{n}$ is just the probability density. Note the dP **is a conditional probability**.

Galaxy field is the non-random field, therefore we can write the conditional probability as:

$$dP(r) = \bar{n}[1 + \xi(r)]dV, \tag{2}$$

where $\xi(r)$ is a factor vs random field, if $\xi(r) > 0$ means clustering, $\xi(r) = 0$ means random field and $\xi(r) < 0$ means void.

The $\xi(r)$ we called 2-point correlation function, this is the initial definition. We also have more direct definition:

$$\xi(r) = \langle \delta(r_0)\delta(r_0 + r) \rangle \qquad \delta(r) = \frac{n(r) - \bar{n}}{\bar{n}}. \tag{3}$$

Why the two beautiful definition is equivalent? Now do a calculation: For non-random point field, we write the conditional probability directly as:

$$\begin{aligned}
dP =& \frac{p(r_0, r + r_0)dV_0 dV}{p(r_0)dV_0} = \frac{p(r_0, r + r_0)}{p(r_0)}dV = \frac{p(r_0, r + r_0)}{\bar{n}}dV \\
p(r_0, r + r_0) &= \langle n(r_0)n(r + r_0) \rangle = \langle \bar{n}[1 + \delta(r_0)]\bar{n}[1 + \delta(r + r_0)] \rangle \\
&= \bar{n}^2 \langle 1 + \delta(r_0) + \delta(r + r_0) + \delta(r_0)\delta(r + r_0) \rangle \\
&= \bar{n}^2 [1 + \langle \delta(r_0)\delta(r + r_0) \rangle] \\
dP =& \bar{n}[1 + \langle \delta(r_0)\delta(r + r_0) \rangle] \qquad \xi(r) = \langle \delta(r_0)\delta(r + r_0) \rangle.
\end{aligned} \tag{4}$$

We introduce another method, confirm the equivalent of two formulas, it will be useful later:

$$\begin{aligned}
\langle\delta(r_0)\delta(r+r_0)\rangle &= \langle\frac{n(r_0)-\bar{n}}{\bar{n}}\frac{n(r_0+r)-\bar{n}}{\bar{n}}\rangle \\
&= \frac{1}{\bar{n}^2}\langle[n(r_0)-\bar{n}][n(r_0+r)-\bar{n}]\rangle \\
\text{important} \qquad &= \frac{\langle n(r_0)n(r_0+r)\rangle - \langle\bar{n}[n(r_0)+n(r_0+r)]\rangle + \langle\bar{n}^2\rangle}{\bar{n}^2} \\
&= \frac{\langle n(r_0)n(r_0+r)\rangle}{\bar{n}^2} - 2 + 1 = \frac{\langle n(r_0)n(r_0+r)\rangle}{\bar{n}^2} - 1 \\
\langle n(r_0)n(r_0+r)\rangle &= \bar{n}^2[1+\langle\delta(r_0)\delta(r+r_0)\rangle] = \bar{n}^2[1+\xi(r)].
\end{aligned} \tag{5}$$

Consequently, we confirm the equivalence by confirming the **joint probability density** $\langle n(r_0)n(r_0+r)\rangle$.

Now we introduce the concept of point pairs: DD, RR and DR. Here D means data and R means random, they are **probability or expected counts** of pairs of target (galaxy here) in the data or random catalogs between the data and random catalogs in given separation bins. We directly write these definitions:

$$\begin{aligned}
DD(r) &= \frac{N_d(N_d-1)}{2}[1+\xi(r)]\frac{4\pi r^2 dr}{V} \\
RR(r) &= \frac{N_r(N_r-1)}{2}\frac{4\pi r^2 dr}{V} \\
DR(r) &= N_d N_r \frac{4\pi r^2 dr}{V},
\end{aligned} \tag{6}$$

where N_d, N_r are the counts of points in given catalogs. Now we can say that $\langle n(r_0)n(r_0+r)\rangle dV = DD(r)$; $\bar{n}^2 dV = RR(r)$ and $\langle\bar{n}n(r)\rangle dV = DR(r)$. Consequently, we get the common estimates of $\xi(r)$ that

$$\begin{aligned}
\xi(r) &= \frac{DD}{RR} - 1 && \textit{Peebles–Hauser} \\
\xi(r) &= \frac{DD - 2DR + RR}{RR} && Landy - Szalay
\end{aligned} \tag{7}$$

the Landy-Szalay estimator comes from Eq. (5) important mark, in practice we usually use the alter quantities:

$$\begin{gathered}
DD' = \frac{DD}{N_d(N_d-1)/2};\ DR' = \frac{DR}{N_d N_r};\ RR' = \frac{RR}{N_r(N_r-1)/2}, \\
\xi(r) = \frac{DD' - 2DR' + RR'}{RR'}.
\end{gathered} \tag{8}$$

Now we extend this theory to continuous fields like density field and 21cm brightness temperature fields, we use density field first because the over-density definition is consistent with over-count number density we defined.

Write these relation directly:

$$
\begin{aligned}
&\xi(r) = \langle \delta_\rho(r_0)\delta_\rho(r+r_0)\rangle \qquad \delta_\rho(r) = \frac{\rho(r)-\bar\rho}{\bar\rho} \\
&\langle \rho(r_0)\rho(r+r_0)\rangle = \bar\rho^2[1+\xi(r)]
\end{aligned} \tag{9}
$$

Now sum up all useful relations and subscripts d and c denote the discrete and continuous field:

$$
\begin{cases}
\xi_d(r) = \langle \delta_d(r_0)\delta_d(r+r_0)\rangle \\
\xi_c(r) = \langle \delta_\rho(r_0)\delta_\rho(r+r_0)\rangle \\
\uparrow \text{Definitions} \\
\langle n(r_0)\, n(r_0+r)\rangle = \bar n^2[1+\xi_d(r)] \\
\langle \rho(r_0)\, \rho(r_0+r)\rangle = \bar\rho^2[1+\xi_c(r)] \\
\uparrow \text{Joint Probability density} \\
dP = \bar n[1+\xi_d(r)]dV \\
\langle \delta_2|\delta_1\rangle = \dfrac{\xi(r)}{\sigma^2}\,\delta_1 \\
\uparrow \text{Conditional Probability}
\end{cases}
\qquad
\begin{aligned}
\xi_d(r) &= \frac{DD' - 2DR' - RR'}{RR'} \\
&= \frac{1}{RR}\Big[DD\frac{N_r(N_r+1)/2}{N_d(N_d-1)/2} - 2DR\frac{N_r(N_r+1)/2}{N_dN_r} - RR\Big].
\end{aligned} \tag{10}
$$

Finally confirm the relation $\langle \delta_2|\delta_1\rangle = \dfrac{\xi(r)}{\sigma^2}\delta_1$:

$$
\begin{cases}
p(\delta_2|\delta_1) = \dfrac{p(\delta_1,\delta_2)}{p(\delta_1)}, \qquad \langle \delta_2|\delta_1\rangle = \displaystyle\int \delta_2\, p(\delta_2|\delta_1)\, d\delta_2 \\
\text{if Gauss,} \quad p(\delta_1,\delta_2) = \dfrac{1}{2\pi\sqrt{\det C}}\ \exp\Big[-\dfrac{1}{2}\,(\delta_1,\delta_2)\, C^{-1}\begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix}\Big] \\
C = \begin{pmatrix}\sigma^2 & \xi(r)\\ \xi(r) & \sigma^2\end{pmatrix} \quad \Rightarrow \quad \langle\delta_1\delta_1\rangle = \langle\delta_2\delta_2\rangle = \sigma^2, \quad \langle\delta_1\delta_2\rangle = \xi(r) \\
p(\delta_1) = \dfrac{1}{\sqrt{2\pi}\,\sigma}\ \exp\Big[-\dfrac{\delta_1^2}{2\sigma^2}\Big] \\
\Rightarrow\ p(\delta_2|\delta_1) = \dfrac{1}{\sqrt{2\pi}}\,\dfrac{1}{\sqrt{\sigma^2-\xi^2/\sigma^2}}\ \exp\Big[-\dfrac{1}{2}\,\dfrac{(\delta_2-\mu_c)^2}{\sigma^2-\xi^2/\sigma^2}\Big], \qquad \mu_c = \dfrac{\xi(r)}{\sigma^2}\,\delta_1 \\
\Rightarrow \quad \langle\delta_2|\delta_1\rangle = \dfrac{\xi(r)}{\sigma^2}\delta_1
\end{cases} \tag{11}
$$

2 Beyond 2PCF

Review:

$$
\xi(r) \equiv \langle \delta(r_0)\,\delta(r_0+r)\rangle, \qquad \delta(r) \equiv \frac{n(r)-\bar n}{\bar n} \quad \text{or} \quad \delta_\rho(r) \equiv \frac{\rho(r)-\bar\rho}{\bar\rho}. \tag{12}
$$

Gaussian random field (GRF) can be fully specified by its mean (zero order) and the 2-point function $\xi(r)$ (or equivalently the power spectrum $P(k)$). Thus, for a GRF, **all** higher moments are determined by ξ via expansion (Wick's theorem below), and all connected $n(\geq 3)$-point cumulants vanish.

For **non-Gaussian** fields, the 2PCF is not sufficient, so we need higher N-point correlation functions to capture non-Gaussian information, e.g., 3-point and 4-point:

$$\langle \delta(r_1)\delta(r_2)\delta(r_3) \rangle, \quad \langle \delta(r_1)\delta(r_2)\delta(r_3)\delta(r_4) \rangle, \ \ldots \tag{13}$$

We define them and show how to expand them into lower order functions (cumulants) below.

Define the n-point correlation function is **moment** of order n is $\langle \delta_1\delta_2 \cdots \delta_n \rangle$. The **connected** n-point function (the n-th cumulant) is denoted $\langle \delta_1 \cdots \delta_n \rangle_c$. They are related by the standard moment–cumulant relations (These relations follow from the cumulants theory detail see wikepedia, simply: generating function $\ln\langle e^{t\delta} \rangle$). For the first few orders (zero-mean fields so that $\langle \delta \rangle = \langle \delta \rangle_c = 0$ for over-density fields). The 2pcf can be expand as below:

$$\langle \delta_1\delta_2 \rangle = \langle \delta_1\delta_2 \rangle_c \equiv \xi(r), \tag{14}$$

we can see it is just 2-th cumulant.

For 3-point correlation function, it also is 3-th cumulant for over-density fields (or all mean is 0 field), we still expand it fully as below:

$$\begin{aligned} \langle \delta_1\delta_2\delta_3 \rangle =& \langle \delta_1\delta_2\delta_3 \rangle_c \\ &+ \langle \delta_1\delta_2 \rangle_c \langle \delta_3 \rangle_c + \langle \delta_1\delta_3 \rangle_c \langle \delta_2 \rangle_c + \langle \delta_3\delta_2 \rangle_c \langle \delta_1 \rangle_c \\ &+ \langle \delta_1 \rangle_c \langle \delta_2 \rangle_c \langle \delta_3 \rangle_c \end{aligned} \tag{15}$$

For 4-point correlation function ignore the terms involve $\langle \delta_i \rangle$, give:

$$\begin{aligned} \langle \delta_1\delta_2\delta_3\delta_4 \rangle &= \langle \delta_1\delta_2\delta_3\delta_4 \rangle_c + \xi_{12}\xi_{34} + \xi_{13}\xi_{24} + \xi_{14}\xi_{23} \\ &= \langle \delta_1\delta_2\delta_3\delta_4 \rangle_c + 3\,\xi^2(r) \qquad \text{if all } r_{ij} = r, \end{aligned} \tag{16}$$

with $\xi_{ij} \equiv \xi(|r_i - r_j|)$.

For 5-point correlation function, emmmmmm, simply:

$$\langle \delta_1\delta_2\delta_3\delta_4\delta_5 \rangle = \langle \delta_1\delta_2\delta_3\delta_4\delta_5 \rangle_c + \zeta_{123}\xi_{45} + \zeta_{124}\xi_{35} + \cdots + \zeta_{345}\xi_{12}, \tag{17}$$

with the symbol ζ_{ijk} denote the 3-th cumulant. Explicitly, $\zeta_{123}\xi_{45} + \zeta_{124}\xi_{35} + \cdots + \zeta_{345}\xi_{12}$ (ten terms), since any partition of $\{1, 2, 3, 4, 5\}$ into a triplet and a pair contributes.

2.1 Wick's theorem (for Gaussian fields)

For a GRF, all cumulants beyond the second order vanish:

$$\langle \delta_1 \delta_2 \cdots \delta_n \rangle_c = 0, \quad n \geq 3. \tag{18}$$

Consequently, any order moment is the sum over all pairings of ξ, and any odd-order moment is zero:

$$\langle \delta_1 \delta_2 \delta_3 \rangle = 0, \qquad \langle \delta_1 \delta_2 \delta_3 \delta_4 \rangle = \xi_{12}\xi_{34} + \xi_{13}\xi_{24} + \xi_{14}\xi_{23}, \ \ldots \tag{19}$$

This is precisely Wick's theorem: **moments can be described by sums of 2-point contractions**.

Finally we explain Why a 3-point function can be nonzero? In a non-Gaussian field the 3-th cumulant $\langle \delta_1 \delta_2 \delta_3 \rangle_c \equiv \zeta$ need not vanish. Writing the full moment,

$$\langle \delta_1 \delta_2 \delta_3 \rangle = \underbrace{\langle \delta_1 \delta_2 \delta_3 \rangle_c}_{\zeta} + \underbrace{\langle \delta_1 \rangle \langle \delta_2 \delta_3 \rangle + 2 \text{ perms}}_{=0 \text{ for zero-mean}}, \tag{20}$$

we can see that the only surviving term is the 3-th cumulant ζ. Therefore, $\zeta \neq 0$ (equivalently, bispectrum $B \neq 0$) is information of non-Gaussianity.

It is often convenient to denote the Fourier transforms of the connected n-point functions: power spectrum P (2-point), bispectrum B (3-point), trispectrum T (4-point), and so on. We will use these to derive "non-standard 2PCFs" later.

2.2 Non-standard 2PCFs that carry higher-order information

Now we derive some two-point statistics that look like 2PCFs but correspond to higher-order information.

2.2.1 $\langle \delta^2(r_0)\delta(r_0 + r)\rangle$ and the bispectrum

Real-space definition

$$\xi_{\delta^2,\delta}(r) \equiv \left\langle \left[\delta(r_0)\right]^2 \delta(r_0 + r) \right\rangle. \tag{21}$$

Fourier representation Define Fourier transforms $\delta(\mathbf{k}) = \int d^3x \, e^{-i\mathbf{k}\cdot\mathbf{x}} \delta(\mathbf{x})$, and $\delta^2(\mathbf{k}) = \int \frac{d^3q}{(2\pi)^3} \, \delta(\mathbf{q})\delta(\mathbf{k} - \mathbf{q})$. The **cross-power** between δ^2 and δ is

$$P_{\delta^2,\delta}(\mathbf{k}) \equiv \langle \delta^2(\mathbf{k}) \, \delta(\mathbf{k}') \rangle = \int \frac{d^3q}{(2\pi)^3} \, \langle \delta(\mathbf{q}) \, \delta(\mathbf{k} - \mathbf{q}) \, \delta(\mathbf{k}') \rangle. \tag{22}$$

The bispectrum B:

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle \equiv (2\pi)^3 \delta_D(\mathbf{k_1} + \mathbf{k_2} + \mathbf{k_3}) B(\mathbf{k_1}, \mathbf{k_2}, \mathbf{k_3}). \tag{23}$$

We can see the third variate $\mathbf{k}'$ must be $-\mathbf{k}$. Therefore,

$$P_{\delta^2,\delta}(k) \equiv \langle \delta^2(\mathbf{k})\, \delta(-\mathbf{k}) \rangle = \int \frac{d^3q}{(2\pi)^3}\, B\big(\mathbf{q}, \mathbf{k}-\mathbf{q}, -\mathbf{k}\big). \tag{24}$$

Thus,

$$\xi_{\delta^2,\delta}(r) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{k}\cdot\mathbf{r}}\, P_{\delta^2,\delta}(k) \quad \Longleftrightarrow \quad \xi_{\delta^2,\delta} \text{ probes } B. \tag{25}$$

For a Gaussian field $B = 0 \Rightarrow P_{\delta^2,\delta} = 0 \Rightarrow \xi_{\delta^2,\delta} = 0$. Nonzero measurement is a clean non-Gaussian signature.

2.2.2 $\langle \delta^2(r_0)\, \delta^2(r_0 + r) \rangle$ and the trispectrum

Define

$$\xi_{\delta^2,\delta^2}(r) \equiv \big\langle [\delta(r_0)]^2\, [\delta(r_0 + r)]^2 \big\rangle. \tag{26}$$

Fourier cross-power:

$$P_{\delta^2,\delta^2}(\mathbf{k}) = \int \frac{d^3q}{(2\pi)^3} \frac{d^3p}{(2\pi)^3}\, \langle \delta(\mathbf{q})\delta(\mathbf{k}-\mathbf{q})\delta(-\mathbf{p})\delta(\mathbf{p}-\mathbf{k}) \rangle. \tag{27}$$

Insert the 4-point expansion (moment = connected + Wick terms):

$$\begin{aligned} P_{\delta^2,\delta^2}(k) = &\int \frac{d^3q}{(2\pi)^3} \frac{d^3p}{(2\pi)^3}\, T(\mathbf{q}, \mathbf{k}-\mathbf{q}, -\mathbf{p}, \mathbf{p}-\mathbf{k}) \\ &+ 2 \int \frac{d^3q}{(2\pi)^3}\, P(q)\, P(|\mathbf{k}-\mathbf{q}|) + \sigma^4, \end{aligned} \tag{28}$$

(here $\sigma^2 \equiv \langle \delta^2 \rangle$). The second term ("$2PP$") is the Gaussian piece (exists even if $T = 0$); the first term is the connected trispectrum contribution. Transforming back:

$$\xi_{\delta^2,\delta^2}(r) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} P_{\delta^2,\delta^2}(k). \tag{29}$$

2.2.3 $\langle \delta^3(r_0)\, \delta(r_0 + r) \rangle$

Define $\xi_{\delta^3,\delta}(r) \equiv \langle [\delta(r_0)]^3\, \delta(r_0 + r) \rangle$. Fourier cross-power becomes a convolution of the **trispectrum** and P plus a product of the bispectrum with $\langle \delta^2 \rangle$:

$$\begin{aligned} P_{\delta^3,\delta}(k) = &\int \frac{d^3q}{(2\pi)^3} \frac{d^3p}{(2\pi)^3}\, T(\mathbf{q}, \mathbf{p}, \mathbf{k}-\mathbf{q}-\mathbf{p}, -\mathbf{k}) \\ &+ 3\, \sigma^2 P(\mathbf{k}). \end{aligned} \tag{30}$$

Again $\xi_{\delta^3,\delta}(r)$ is its Hankel transform (detail see Appendix).

Why "3" factors appear? They count the number of ways to choose which δ inside δ^3 participates in the connected contraction.

2.3 Marked Correlation Function

Another method to probe non-Gaussian fields is marked correlation function (MCF), we write the original formula as below:

$$\langle m(r_0)m(r_0+r)\rangle = \frac{1}{n_{\text{pair}}(r)} \sum m_i m_j \,, \tag{31}$$

where m is the marked value; therefore, m_i is the marked value at a given point and $\bar{m}$ is the space average of the total marked value, $n_{\text{pair}}(r)$ is the total counts of pair pairs of points at distance r which means the average. If mark is continuous, consider a point with mark larger than this mean value. Are the point neighbouring it also likely to have larger marks? One way to quantify this is to compute the ratio of the mean mark to $\bar{m}$ of pairs of point as a function of pair separation like Beisbart2000's $k_{mm}(r)$ which is $\langle \frac{m(r_0)}{\bar{m}} \frac{m(r_0+r)}{\bar{m}} \rangle$, we directly write the important relation:

$$M(r) = \frac{1}{n_{\text{pair}}(r)\bar{m}^2} \sum m_i m_j = \frac{1+W(r)}{1+\xi(r)}, \tag{32}$$

here $M(r)$ called MCF. We can see that MCF is expressed as the ratio of a new correlation function $W(r)$ to 2PCF. So we can easily find the meaning of mcf is just like 2pcf in an inhomogeneous field vs. a random field, and it says that the mark you chosen corelation vs. the 2pcf. For example, $M(r) > 1$ describes the clustering of marks larger than the clustering itself.

Now we derive the MCF. Our target is why MCF can be expressed as a ratio of a new correlation function $W(r)$ to 2PCF and what is $W(r)$.

We can treat MCF and the mark product average as conditional expectation:

$$\begin{aligned} n_{\text{pair}}(r) &= \int \langle n(r_0)n(r_0+r)\rangle \, dV_1 dV_2 \\ \sum m_i m_j &= \int \langle m(r_0)n(r_0)\, m(r_0+r)n(r_0+r)\rangle \, dV_1 dV_2 \\ p(m(r_0), m(r_0+r)|r) &= \frac{\langle m(r_0)n(r_0)\, m(r_0+r)n(r_0+r)\rangle}{\langle n(r_0)n(r_0+r)\rangle} \\ \langle m(r_0)m(r_0+r)\rangle &= \int dm_1 dm_2 \, m_1 m_2 \, p(m(r_0), m(r_0+r)|r), \end{aligned} \tag{33}$$

define the two-point function of the weighted point process:

$$1+W(r) = \frac{\langle m(r_0)n(r_0)\, m(r_0+r)n(r_0+r)\rangle}{\bar{m}^2\bar{n}^2}, \tag{34}$$

we get:

$$M(r) = \frac{1+W(r)}{1+\xi(r)}. \tag{35}$$

2.3.1 Expansion

Consider a continuous field $\rho(\mathbf{r})$ here is density field, treat $n(r) = \bar{n} = 1$ at any r:

$$\begin{aligned}1 + W_\rho(r) &= \frac{\langle \rho(\mathbf{r_0})\rho(\mathbf{r_0} + \mathbf{r})\rangle}{\bar{\rho}^2} \\ &= \frac{\langle \bar{\rho}(1 + \delta(\mathbf{r_0}))\bar{\rho}(1 + \delta(\mathbf{r_0} + \mathbf{r}))\rangle}{\bar{\rho}^2} \\ &= \langle (1 + \delta(\mathbf{r_0}))(1 + \delta(\mathbf{r_0} + \mathbf{r}))\rangle \\ &= 1 + \langle \delta(\mathbf{r_0})\delta(\mathbf{r_0} + \mathbf{r})\rangle \\ &= 1 + \xi_\delta(r)\,,\end{aligned} \tag{36}$$

because the point distribution is fully uniformity, we have

$$1 + \xi(r) = 1, \tag{37}$$

therefore, we get

$$M_\rho(r) = \frac{1 + W_\rho(r)}{1 + \xi(r)} = 1 + \xi_\delta(r). \tag{38}$$

We found for the density continuous fields the 2PCF can be treated as that mark is density itself MCF, because the point process in continuous field becomes 1.

2.3.2 General for continuous fields

The uniform-sampling result above corresponds to taking the baseline sampling intensity to be constant. More generally, let $I(\mathbf{r})$ be a indicator that determines how pairs are considered, and let $m(\mathbf{r})$ be a mark field with $\bar{m} \equiv \langle m\rangle$. Define

$$\begin{aligned}1 + \xi_I(r) &\equiv \Big\langle \frac{I(\mathbf{r}_0)}{\bar{I}}\,\frac{I(\mathbf{r}_0 + \mathbf{r})}{\bar{I}}\Big\rangle, \\ 1 + W_{I,m}(r) &\equiv \Big\langle \frac{I(\mathbf{r}_0)}{\bar{I}}\,\frac{I(\mathbf{r}_0 + \mathbf{r})}{\bar{I}}\,\frac{m(\mathbf{r}_0)}{\bar{m}}\,\frac{m(\mathbf{r}_0 + \mathbf{r})}{\bar{m}}\Big\rangle,\end{aligned} \tag{39}$$

so we get

$$M_I(r) \equiv \frac{1 + W_{I,m}(r)}{1 + \xi_I(r)} = \frac{\langle m(\mathbf{r}_0)\, m(\mathbf{r}_0 + \mathbf{r})\rangle_{I\text{-weighted}}}{\bar{m}^2}, \tag{40}$$

where $\langle\cdot\rangle_{I\text{-weighted}}$ denotes the conditional average over all pairs at separation r with pair weight $I(\mathbf{r}_0)I(\mathbf{r}_0 + \mathbf{r})$.

Two classification.

1. Uniform (continuous approach $I \equiv 1$) $1 + \xi_I(r) = 1$ and

$$M(r) = 1 + W_{I,m}(r) = \Big\langle \frac{m(\mathbf{r}_0)}{\bar{m}}\,\frac{m(\mathbf{r}_0 + \mathbf{r})}{\bar{m}}\Big\rangle = 1 + \xi_m(r). \tag{41}$$

This reproduces the density example above when choosing $m = \rho/\bar{\rho} = 1+\delta$.

2. Mass/intensity weighting (I traces a physical intensity).

For an "number density" view of continuous fields, we consider that the "number density" for fields is just the volume density of physical quantities.

For example, choosing $I = \rho$ (mass weighting) in density fields and mark is any other functions of density we have

$$\begin{gathered} 1+\xi_I(r) = \Big\langle \frac{\rho}{\bar{\rho}} \frac{\rho'}{\bar{\rho}} \Big\rangle, \quad 1+W_{I,m}(r) = \Big\langle \frac{\rho}{\bar{\rho}} \frac{\rho'}{\bar{\rho}} \frac{m}{\bar{m}} \frac{m'}{\bar{m}} \Big\rangle = \frac{1}{\bar{\rho}^2 \bar{m}^2} \langle \rho\rho' m m' \rangle, \\ M_I(r) = \frac{1+W_{I,m}(r)}{1+\xi_I(r)}. \end{gathered} \tag{42}$$

Note that, unless $I \equiv \text{const}$, the numerator is **not** simply $1+\xi_m(r)$ but the two-point function of the composite field $(I/\bar{I})\,(m/\bar{m})$.

For 21cm coevals, the redshifts is fixed at given coeval, so the frequency is fixed, we can consider the photon number density to be proportional to the brightness temperature $I_T(\mathbf{r}) \propto T_b(\mathbf{r})$, and let the mark be $m(\mathbf{r}) = f\big(T_b(\mathbf{r})\big)$.

Note the 21cm brightness temperature 2PCF defined by $\langle [T(\mathbf{r}_0) - \bar{T}][T(\mathbf{r}_0 + \mathbf{r}) - \bar{T}]\rangle\rangle$. Therefore

$$\begin{aligned} \xi_T(r) &\equiv \langle [T(\mathbf{r}_0) - \bar{T}][T(\mathbf{r}_0+\mathbf{r}) - \bar{T}]\rangle\rangle \\ 1+W_{T,f}(r) &\equiv \Big\langle \frac{T(\mathbf{r}_0)}{\bar{T}} \frac{T(\mathbf{r}_0+\mathbf{r})}{\bar{T}} \left[f(T_b(\mathbf{r}_0)) - \bar{f}\right] \left[f(T_b(\mathbf{r}_0+r)) - \bar{f}\right] \Big\rangle, \\ M_T(r) &= \frac{1+W_{T,f}(r)}{\Big\langle \frac{T(\mathbf{r}_0)}{\bar{T}} \frac{T(\mathbf{r}_0+\mathbf{r})}{\bar{T}} \Big\rangle} = \frac{1+W_{T,f}(r)}{1+\xi_T(\mathbf{r})/\bar{T}^2} \end{aligned} \tag{43}$$

I'm not ensure the handle of 21cm field is true.

Remarks. When m is a nonlinear function of the underlying field (e.g., $m = f(\delta)$ or $m = f(T_b)$), the expansion of $1 + W_{I,m}$ generally involves higher-order statistics (bispectrum/trispectrum), which explains the MCF sensitivity to non-Gaussianity.