Statistics of merging peaks of random Gaussian fluctuations: skeleton tree formalism

Hitoshi Hanami

ESO/Santiago, Casilla 19001, Santiago 19, Chile

ABSTRACT
Cosmological bound objects are generally considered to be formed from the local maxima of cosmological density fluctuations, which are often assumed to be Gaussian random fields. In order to study the statistics of those objects that are the result of hierarchical merging, we propose skeleton tree formalism, which can analytically distinguish episodic merging from continuous accretion in the mass-growth processes. This distinction was not clear in the extended Press–Schechter (PS) formalism. Skeleton tree formalism is a natural extension of peak theory, which is an alternative formalism for the statistics of bound objects. Fluctuation-field smoothing with a Gaussian filter produces a landscape by adding the extra dimension of the filter-resolution scale to the spatial coordinate of the fluctuation. In this landscape, a smoothed peak is nesting alongside the neighbouring peak and appears as a critical point called a sloping saddle, which can be interpreted as a destroyed object during a merger event in the context of the hierarchical structure formation. The topological properties of the landscape can be abstracted as skeleton trees, which consist of the line process of the smoothing peaks and the point process of the sloping saddles. According to this abstract topological picture, in this paper, we present the concept and the basic results of skeleton tree formalism in describing (1) the distinction between accretion and a merger in the hierarchical structure formation from various initial random Gaussian fields; (2) the instantaneous number density of the sloping saddles, which gives the destruction rate in the mergers; (3) the number densities of all the peaks, nesting and non-nesting, with simple and analytical distinguishing definitions; (4) the self-consistency of the formalism in reproducing the statistics, including the conservation equation, of all the peaks; (5) the instantaneous number density of the reforming objects, which is derived from the conservation formula of the non-nesting peak; and (6) the mean growth history of the objects, which is reproduced from the rates of destruction, reformation, and from the relative accretion growth.

Key words: galaxies: clusters: general – galaxies: formation – cosmology: theory – dark matter.

1 INTRODUCTION
The hierarchical clustering scenario, including the cold dark matter (CDM) model, may be the most established one for reconstructing various observational properties in cosmological structure from galaxies to clusters of galaxies. Press & Schechter (1974, hereafter PS) first proposed an analytical formalism which derived the number density of bound virialized objects with a mass at any given epoch, with the assumption that the primordial density fluctuations were random Gaussian fields. The mass function predicted by the PS theory shows reasonably good agreement with N-body simulations even if it has more low-mass objects (e.g. Lacey & Cole 1994). The original PS theory was extended by Bower (1991) and Bond et al. (1991) in order to derive the number density of objects with a certain mass at a given time subject to a larger object at a later time. Using the formalism, Lacey & Cole (1993, hereafter LC) reconstructed the hierarchical mass growth of cosmic virialized haloes known as ‘merger tree’.

The PS formalism, however, has a limitation when used to describe the history of the mass growth of individual objects. The ‘merging process’ described using the PS approach in LC cannot be...
interpreted as having the same meaning as a ‘merger’ in the astronomical sense, since objects lose their identity in the merging process. In the mass-growth history of astronomical objects, continuous accretion on to a bound object without the loss of identity has a different meaning from mass accumulation (the loss of the identity) in a major merger. The formalism based on the PS approach cannot imply the distinction between ‘tiny’ and ‘notable’ captures.

For solving this problem, Manrique & Salvador-Sole (1995, 1996; hereafter MS95 and MS96) and their collaborators (Salvador-Sole et al. 1999) proposed a formalism which they called the ‘Confluent System of Peak trajectories’ (CUSP) formalism as an extension of adaptive windowing by Appel & Jones (1990). This can be categorized as a type of ‘peak’ theory (Doroshkevich 1970; Adler 1981; Bardeen et al. 1986, hereafter BBKS), which can count the number of density peaks related to the collapsing threshold, applying a low-pass filter of the bound-object scale to the fluctuation field. While the CUSP formalism can describe the merger as distinct from accretion in an analytical way, unfortunately, it is not easy to get a simple picture of the merger from the formalism directly, since the rates of destruction and reformation in the merger can be represented only by the integral form of the joint number density of nesting peaks in the background peak. This calculation cost also makes it difficult to apply the scheme to semi-analytical studies of galaxy formation, including the mass accumulation history of bound objects which is studied with the Monte Carlo method in LC.

In order to give an analytical description of a merger using a simple picture of adaptive windowing, we developed a new approach called skeleton tree formalism, using topological characteristics to smoothen the random field with a Gaussian filter. Picking up sloping saddles in the landscape of the smoothed field, we can identify destroyed objects in the mergers. The topological features in the landscape that we call ‘sloping saddles’ can be reduced into skeleton trees. This tree picture gives a clear illustration of nesting in the mergers, from which we can derive the number density of the nesting and the non-nesting peaks with simple and analytical definitions.

In this paper, we will focus on the concept and the basic description of skeleton tree formalism. The outline of the paper is as follows. In Section 2, in order to distinguish the merger process from the accretion process in our context, we sketch the topological characteristics in the landscape of the smoothed field, picking up the critical points of the peaks and the sloping saddles. The critical points and the smoothing process in the field reconstruct the formation and growth history of the objects with a skeleton tree picture. In Sections 3, 4, and 5, we formulate the mathematical description of the constraints, the probability distribution functions, and the scale-functions for the critical points. In Section 5, the mass function is also derived with a reasonable mass-assignment scheme. In Section 6, we show the results of the evolution rates with regard to accretion, destruction and reformation in the merger obtained from skeleton tree formalism, with its high rate of consistency. Finally, we present our conclusions in Section 7. We have relegated the details of the derivations to four appendices.

2 IDENTIFICATION OF ACCRETION AND MERGING

2.1 Hierarchical evolution from fluctuations and the filtering process

We will express the density fluctuation field as the functions of the comoving spatial coordinates \( r \) and \( k \):

\[
\delta(r) = \int d^3k e^{ikr} \delta(k).
\]

Interesting collapse objects with a comoving scale \( R \) can be identified as peaks higher than a given collapse threshold in the fields smoothed with a low-pass filter of the resolution scale \( R \). The fluctuation, smoothed with the selection function \( S(r;R) \), can be expressed as

\[
F(r;R) = \int d^3r_0 S(r_0;R) \delta(r - r_0) \int d^3r_0 S(r_0;R) = 1.
\]

This Fourier transform is represented with a window function which is the Fourier transform of the selection function,

\[
F(k;R) = \delta(k) W(k;R).
\]

In our interesting cases, the window function works as a low-pass filter.

We shall restrict ourselves to isotropic homogeneous Gaussian random fields with zero mean as descriptions of the initial fluctuations. For the field, the power spectrum is then only a function of \( k = |k|; \delta(k) = \delta(k) |k| \). The fluctuation spectrum filtered with the scale \( R \) is

\[
P(k;R) = |k|^2 W(k;R)^2.
\]

In this paper, we take a normalized isotropic Gaussian filter,

\[
W(k;R) = W(k;R) = e^{-k^2R^2/2}.
\]

In the next subsection, we will discuss the reason behind its unique property of monotonic decreasing of the peak height with smoothing.

The linear theory of gravitational instability for structure formation, the amplitude of the field in the overdensity area first grows in proportion to \( D(t) \), where \( D(t) \) is the linear growth factor. According to BBKS, a bound object collapses from the area of a comoving scale \( R \) when the density of a peak in the fluctuation smoothed over the resolution scale \( R \) exceeds a fixed threshold \( \delta_{\text{c};0} \). Instead of viewing the peaks as growing in density amplitude relative to a fixed threshold \( \delta_{\text{c};0} \), we can interpret that the threshold level \( \delta_{\text{c}} \) is decreasing as \( \delta_{\text{c};0} D(t)^{-1} \), fixing the initial fluctuation field as \( F(r;R) \), where \( \delta_{\text{c};0} \) was determined from the threshold at the present time \( t_0 \). Then, the decreasing threshold level can be identified as the time under the monotonic mapping of \( \delta_{\text{c}} = \delta_{\text{c};0} D(t)^{-1} \). In this paper, we take the Einstein–de Sitter model: \( \Omega_0 = 1, \lambda_0 = 0 \), in which the relative threshold level \( \delta_{\text{c}} / \delta_{\text{c};0} = D(t_0) D(t)^{-1} = (1 + z) \), where \( z \) is the redshift at time \( t \).

For standard initial fluctuations like CDM models, in general, the rms of the smoothed field \( (F(r;R)^2) \) decreases with increasing \( R \). For such a fluctuation, which decreases the collapse threshold with time evolution, we can pick the collapse objects in the larger scale, which gives a reasonable sketch of the hierarchical clustering picture. We will relate the filtering process to the hierarchical clustering description in the next subsection.

2.2 Landscape of the fluctuation field in position and resolution space

Consider that the random field in one-dimensional (1D) positional coordinate of \( x \) is smoothed with a low-pass filter of a resolution scale \( R \). This field is reproduced as a landscape figured in a two-dimensional (2D) extended space of \((x;R)\). A smoothing peak with
2.3 Definition of monotonic accretion and merging in the landscape

In order to understand the monotonic evolution of the field smoothed with a Gaussian filter of the resolution scale $R$, we rewrite the derivative of the field as

$$\frac{\partial F(r;R)}{\partial R} = R^{\frac{3}{2}} F(r;R).$$

(6)

This is identical to a diffusion equation of the variables $(R^2, r)$. For all critical points of $\nabla^2 F < 0$ ($\nabla^2 F > 0$) like peaks (holes), it guarantees the monotonically decreasing (increasing) as the scale increasing as

$$\frac{\partial F(r;R)}{\partial R} < 0, \quad \left(\frac{\partial F(r;R)}{\partial R} > 0\right).$$

(7)

This monotonicity of the peak smoothing also guarantees that the peak runs continuously on the ridge to the shore cape without the appearance of an island in the landscape. It is reasonable for the smoothing and merging of peaks to be defined as the continuous accretion growth and the merging events of bound objects. If we have islands in the landscape with filters other than Gaussian, however, we cannot distinguish merging from accretion without there being confusion as to the scale identification of related bound objects. Fortunately, we can exclude this problem as long as we use the Gaussian filter, which guarantees the absence of ‘islands’ in the landscape as shown above. This is the reason why we use a Gaussian filter in this paper.

A ridge in the landscape, then, represents the continuous accretion growth of a bound object. On the other hand, some ridges terminate on the slopes of neighbouring ridges. The vanishing point of the ridge on the slope of the neighbouring ridge can be defined as a type of critical point. We call it a sloping saddle, since it is a saddle point on the slope of the neighbouring peak. The sloping saddle can represent the reasonable feature of a bound object that loses its identity in the merging process, associated with a tree structure in which the branches of the ridges are nested at the joints of the sloping saddles. Then, we can translate the topology in the landscape to the tree structure that we call a skeleton tree, which consists of the accretion branches and the merger joint picked up with the sloping saddles.

2.4 Skeleton tree picture

We will describe the steps by which we abstract from the landscape...
to the tree structure, as presented schematically in Fig. 1. In
Fig. 1(a), the local structure of the landscape is represented by
contour lines, and the different classes of critical points are
marked: the peak is $\bigodot$, the sloping saddle is $\bigcirc$, and the
ridges are represented as dashed lines. In this example, the sloping
saddle appears on a resolution scale $R_c$ at a threshold $\delta_c$. A ridge
terminates at the sloping saddle $\bigodot(R_c,\delta_c)$. It means that a bound
object of the scale $R_c$ loses its identity at threshold $\delta_c$. There is a
ridge neighbouring the sloping saddle, on which the peak has a
density of $\delta_c$ at the same resolution scale of $\bigodot(R_c,\delta_c)$, so we call $\delta_c$
the foreground density. The peak on the neighbouring ridge
reforms around the sloping saddle and becomes an isolated peak
$\bigodot(R_c,\delta_c)$ at the background density $\delta_c$. Then, we can consider that
the reformation starts from $(R_c,\delta_c)$ of a peak neighbouring the
sloping saddle, and ends at $(R_c,\delta_c)$. The merger, with destruction
and reformation, starts from $\delta_c$ and ends at $\delta_c$ with the scale range
from $R_c$ to $R_l$ in the case of Fig. 1(a).

As shown in Fig. 1(b) as an abstracting step of Fig. 1(a), then, we
can introduce a tree structure associated with the sloping saddle
$\bigodot(R_c,\delta_c)$, two vanishing peaks $\bigodot(R_c,\delta_c)$ and $\bigodot(R_c,\delta_c)$ at
the foreground density of $\delta_c$, and a reforming peak $\bigodot(R_c,\delta_c)$ at
the background density of $\delta_c$. The two peaks $\bigodot(R_c,\delta_c)$ and $\bigodot(R_c,\delta_c)$ are
on the ridges of $\bigodot(R_c,\delta_c)$ and $\bigodot(R_c,\delta_c)$ respectively.

To describe the history of hierarchical clustering with merging, it
is natural to reorder the tree structure along the threshold level
instead of on the resolution scale. Fig. 1(c) schematically shows
this reconstructed feature from Fig. 1(b). The continuous accretion
growth is represented as dashed lines along the ridges, while the
merger is represented as the hatched area in Fig. 1(c).

In this case, the merging occurs during the interval between $\delta_c$
and $\delta_c$. Even if we can follow the neighbouring peak along the
same ridge over $\delta_c$ in the landscape, however, we have two nesting
peaks taking part in the merger as a result of the ridge being folded
and reordered around the sloping saddle, as shown in the hatched
area in Fig. 1(c). It means that the progenitor of the reforming peak
on the ridge neighbouring the sloping saddle also loses its identity
in the merger. Then, the neighbouring ridge is divided into two
stages at $\delta_c$. One of them is the ridge that has terminated as the
disrupting progenitor $\geq \delta_c$, and the other is the ridge that started as a
reforming peak $< \delta_c$. The merger feature is simplified as a joint,
which connects the three lines of (1) a disappearing peak that has
terminated at a sloping saddle and (2) two nesting peaks (one a
progenitor, one a reformer) around the point of the sloping saddle,
where the lines are the abstraction of the ridges representing
continuous accretion growth. We note that the branching number
on the progenitor side is always two for a merger. This is natural,
since the probability of mergers of three or more multi-progenitor
peaks are negligible as compared the probability of mergers of two
progenitor peaks. Then, we can practically reproduce a merger
with a destroyed peak and two nesting peaks (a progenitor and a
reformer), as three line processes of the accretion branches are
connected around a sloping saddle as a point process of $(R_c,\delta_c)$
marked as $\bigotimes$ as shown in Fig. 1(d), which makes a skeleton tree.
These lines and point processes can be given as mathematical
definitions as introduced in the next section.

3 THE ENSEMBLE AVERAGED DENSITIES
OF CRITICAL POINTS
In this section, we will describe the mathematical representation of
the constraints which pick up peaks and sloping saddles as line and
point processes in the skeleton tree formalism. With constraints,
we give the general formula of the ensemble averaged density.

3.1 Line and point processes in the skeleton tree
As shown in the previous section, the skeleton tree formalism is
described by a set of the line and point processes of the smoothing
peak along the ridge and the sloping saddle in the extended space
$(r, R)$. The line process of the smoothing peaks in $(r, R)$ is equal to
the point process in the original spatial coordinate $r$.

The density fields of these point processes are described as sums of $\delta$
functions,
\begin{equation}
n_{pk}(r,R) = \sum_i \delta^{(3)}(r - r_{pk,i}),
\end{equation}
\begin{equation}
n_{ss}(r,R) = \sum_i \delta^{(3)}(r - r_{ss,i})\delta(R - R_{ss,i}),
\end{equation}
where the subscripts $pk$ and $ss$ mean means peaks and sloping saddles, and
$R$ is the resolution scale of the sloping saddles. Since $n_{pk} d^3 r$
and $n_{ss} d^3 r dR$ are the numbers in 3D infinitesimal volume of $d^3 r$
and 4D infinitesimal volume of $d^3 r dR$, we treat them the spatial density
and the instantaneous spatial density, respectively.

We will express the point processes entirely in terms of the field
and its derivatives with the spatial coordinate $r$ and the resolution
scale $R$. In the neighbourhood around a critical point, with its
constraint of $\nabla F(r)|_{cr} = 0$, we can expand the field in a Taylor
series,
\begin{equation}
F(r) = F(r_{cr}) + \frac{1}{2!} \nabla \nabla F(r_{cr}) \Delta r \Delta r + \frac{1}{3!} \nabla \nabla \nabla F(r_{cr}) \Delta r \Delta r \Delta r,
\end{equation}
and its derivatives can be also expanded as
\begin{equation}
\nabla F(r) = \nabla \nabla F(r_{cr}) \Delta r,
\end{equation}
\begin{equation}
\nabla \nabla \nabla F(r) = \nabla \nabla \nabla F(r_{cr}) \Delta r + \nabla \nabla \nabla \nabla F(r_{cr}) \Delta r,
\end{equation}
where $\Delta r = r - r_{cr}$ and the subscript $cr$ means the value at the
critical point.

The critical points can be divided into non-degenerate and
generate. The extrema like a peak and hole can be categorized into
a non-degenerate critical point. Provided that the condition of
the non-degenerate extrema $\det(\nabla \nabla \nabla F)|_{cr} \neq 0$, equation (11) can be rewritten as
\begin{equation}
r - r_{cr} = (\nabla \nabla \nabla F(r_{cr})^{-1})^{-1} \nabla F(r).
\end{equation}
Using the second derivatives of the field, the number density of the
extrema can be represented as
\begin{equation}
n_{ex} = \delta^{(3)}(\nabla \nabla \nabla F)^{-1} \nabla F = |\det(\nabla \nabla \nabla F)| \delta^{(3)}(\nabla F).
\end{equation}
In order to describe the point process for the extrema, therefore, it
is enough to take the terms to the order of the second derivatives.

In the degenerate case of $\nabla \nabla \nabla F(r_{cr}) = 0$, however, we cannot
describe the displacement vector with only the first and second
derivatives of the field as the non-degenerate case of equation (13).
The sloping saddle is a type of degenerate critical point.

Under the transformation of the principal axis in the spatial
coordinate, those parts of the six components that are related to the
second derivatives $\nabla \nabla \nabla F(r_{cr})$ in the covariance matrix become
diagonal. It means that the second derivatives have three
eigenvalues.
(\(F_{11}, F_{22}, F_{33}\)) = -\(\sigma_2(\lambda_1, \lambda_2, \lambda_3)\), \(F_{ab} = 0 \) \((\beta \neq \alpha)\),

where \(\sigma_2\) is the rms of \(\nabla \otimes \nabla F\) (see the definition in Appendix A). All the eigenvalues are not null in the non-degenerate cases. On the other hand, the degenerate critical points have at least one null eigenvalue. Therefore we assumed \(\lambda_1 \geq \lambda_2 \geq \lambda_3 = 0\) for convenience. This is the reason for the break of the non-degenerate condition as \(\det(\nabla \otimes \nabla F)_{\text{cr}} = 0\) in the degenerate case. In general, a sloping saddle has a neighbouring peak in the degenerate direction.

In this degenerate direction, we cannot take the inversion of equation (11) as \(x_3 - x_{3,ss} = F_{33}^{-1}F_3\) since \(F_{33}|_{\text{cr}} = 0\).

In order to describe the point process of the sloping saddles, we use the expansions of \(F_{33}\) at the degenerate direction and \(\nabla \otimes \nabla F\) for the rest of the 2D non-degenerate components in a combination of equations (11) and (12),

\[
F_{33} = \sum_{\beta = 1}^{3} F_{33,\beta} \Delta x_{\beta} = F_{333}\Delta x_3,
\]

\[
\nabla^{(2)} F = (\nabla^{(2)} \otimes \nabla^{(2)} F) \Delta x^{(2)} F(r),
\]

where the superscript \((2)\) means the 2D non-degenerate space. The set of the expansions derives as

\[
r^{(2)} - r_{\text{cr}}^{(2)} = (\nabla^{(2)} \otimes \nabla^{(2)} F)_{\text{cr}}^{-1}(\nabla^{(2)} F(r)),
\]

\[
x_3 - x_{3,ss} = F_{333}^{-1}F_{33}.
\]

Furthermore, we should remember that the sloping saddle is defined in extended space by the resolution scale \(R\). The gradient \(\nabla F(r)\) can be expanded with the derivative of \(R\)

\[
\nabla F(r) = \frac{d\nabla F}{dR} |_{\text{ss}} (R - R_{ss}) = \nabla^{2} F |_{\text{ss}} (R - R_{ss}).
\]

In the degenerate direction, this gives

\[
R - R_{ss} = \left( \sum_{\alpha = 1}^{3} F_{33a}R \right)^{-1} F_{33}.
\]

We obtain the constraint for the instantaneous spatial density of the sloping saddles as

\[
n_{ss} = \sum_{i} \delta^{(3)}(r - r_{ss}) |_{\text{ss}} (R - R_{ss}) = \lvert \det(\nabla^{(2)} \otimes \nabla^{(2)} F)\rvert |_{\text{ss}} (\nabla^{(2)} F)
\]

\[
\times |F_{333}| \delta(F_{33}) \sum_{\alpha = 1}^{3} F_{33a}R \rvert \delta(F_{33}).
\]

3.2 Ensemble averaged density of critical points

The joint probability distribution of \(n\)-dimensional random variables with multivariate Gaussians can be described as

\[
p(F^{(n)}) dF^{(n)} = \exp\left[\frac{1}{2} (F^{(n)})^T C^{(n)} (F^{(n)})^{-1} F^{(n)}\right] \frac{1}{(2\pi)^{n/2} \det C^{(n)}} dF^{(n)}.
\]

The covariance matrix \(C^{(n)}\) is defined as the expectation value of the direct product of the vector \(F^{(n)}\):

\[
C_{\alpha \beta}^{(n)} = \left( F_{\alpha}^{(n)} F_{\beta}^{(n)} \right),
\]

where the superscript \((n)\) of \(F^{(n)}\) is the dimension of parameter space.

In general, the conditional probability of the event \(A\) with the constraint event \(B\) is given by the Bayes formula \(p(AB) = p(A|B)p(B)\). The joint probability, then, can be expanded with the divided conditional probabilities. If the parameters in the vector are all Gaussian distributed, we can directly obtain the covariance matrix of the divided conditional probability. There is a general theorem which is extremely useful when we calculate the above joint probability from the divided conditional probabilities. For the count of the peaks, we need only the 10-dimensional parameters \(F^{(10)} = (F, \nabla F, \nabla \otimes \nabla F)\). For the count of the sloping saddles, we should extend the parameter space to the 20-dimensional parameter space \(F^{(20)} = (F, \nabla F, \nabla \otimes \nabla F, \nabla \otimes \nabla \otimes \nabla F)\). The 20-dimensional covariance matrix can be found with the explicit forms of the parameters in Appendix A. With the help of the theorem in Appendix B, the divided conditional probabilities are described for the cases of the peaks and the sloping saddles in Appendices B2 and B3, respectively.

For a type of critical point in an \(n\)-dimensional parameter space described with constraints, in general, the probability weighted density is

\[
n_{ss}[F^{(n)}] dF^{(n)} = p[F^{(n)}]|_{C}(F^{(n)}|\text{critical points}) dF^{(n)}.
\]

Then, the ensemble average density of the type of critical points with \(m\) restrictions can be obtained from the integration over \((n-m)\) dimensional parameters;

\[
\left< n_{ss}[F^{(m)}] \right> dF^{(m)} = \int dF^{(n-m)} n_{ss}[F^{(n)}] dF^{(m)}.
\]

4 SCALE FUNCTIONS OF VARIOUS PEAKS

With the formula obtained in the previous section, we can calculate the ensemble average densities of sloping saddles and peaks by counting the point and line processes. As discussed in Section 2, a sloping saddle can be interpreted as a destroyed peak in a merger. Then, we can introduce the instantaneous scale-function of the destroyed peaks, which is identical to the ensemble number density of the sloping saddles in an infinitesimal region of \(dR \delta_{ss}\). It will be given in the first subsection below.

On the other hand, we should be cautious about introducing the scale-function of peaks from their ensemble number density in an infinitesimal interval of \(dR\) since there are two types of peaks; non-nesting peaks and nesting peaks. We should distinguish the non-nesting peaks from the nesting peaks. With this distinction, from their ensemble number density in an infinitesimal interval of \(dR\), we will also derive their scale-functions.

4.1 Instantaneous scale function of destroyed peaks in mergers

A destroyed peak at a merger event can be identified as a sloping saddle as seen in the skeleton tree. Then, the instantaneous scale-function of destroyed peaks can be directly calculated as the ensemble-averaged density of sloping saddles in the infinitesimal volume \(dR d\nu\) in \((R, \nu)\) space:

\[
N_{ss}(R, \nu) dR d\nu = \langle n_{ss}(R, \nu) \rangle dR d\nu,
\]

where \(\nu = \delta_{ss}(t)/\sigma_0\). The brief calculation is described in Appendix C.

With the transformation from \(\nu\) to \(\delta_{ss}\), the instantaneous scale functions of the disappearing peaks with the scale \(R\) at \(\delta_{ss}\) are
The instantaneous scale-function is responsible for the time-resolution scale in the range of the resolution scale between $\delta R$.

The merger feature with nesting peaks in the skeleton tree formalism. The naive merging picture in the space $(\nu, R)$ is represented in the same way as Fig. 1. We have schematically shown the two cases, (b) and (c), as those without or with nesting as the typical feature of a merger in the space $(\delta R)$. If there are the nesting peaks, the scale jump appears even when the density threshold is fixed, as in (c).

4.2 Scale functions of all the peaks

We introduce the scale-functions of all the peaks $N_{pk}(R, \delta_0) \, dR \, d\delta_0 = N_{\nu_0}(R, \delta_0) \, dR \, d\nu$. The instantaneous scale-function is responsible for the time-evolution of peaks in the merging process, as seen in the skeleton tree of Fig. 1; the line of the smoothing peak connects with other lines at the joint of the sloping saddle.

4.3 Scale function of nesting peaks

As we discussed with the skeleton tree, a set of nesting peaks appears as a pair, consisting of a progenitor and a reformer, in a merger. We will now consider the feature of the nesting peaks with the help of the graphical representation in Fig. 2. We have schematically shown two cases with and without nesting in Fig. 2. If we have two nesting peaks, a progenitor and a reformer, with a scale difference $\Delta R$ at the same density threshold $\delta$ as shown in Fig. 2 (c), we can identify the case with nesting. The scale jump $\Delta R$ induces the displacement of the density contrast $\Delta \nu$ even when the density threshold is fixed; $\delta_0 = $ constant. Then, with the mapping relation in the nesting process between $\nu$ and $R$,

$$ d\nu = \left. \frac{\partial \nu}{\partial R} \right|_{\delta_0=\text{constant}} dR = \nu \left. \frac{d \ln \sigma_0}{d \sigma_0} \right| dR. $$

Remarkably, $N_{pk}(R, \delta) \, d\nu$ is given as the ensemble-averaged density of the peaks in the infinitesimal interval of $\nu$, we can derive the scale-function of the nesting peaks as

$$ N_{pk}^{\nu}(R, \delta) = N_{pk}(R, \delta) \nu_0 \frac{\sigma_0(R)}{\sigma_0} dR, $$

where $\nu_0 = \gamma \nu$. This formula is the same expression as the scale-function of all the peaks initially proposed by Bond (1989). Though the two expressions of $N_{pk}^{\nu}(R, \delta) \, d\nu$ and $N_{pk}^{\nu_0}(R, \delta) \, dR$ are similar to each other, they differ by a factor of $1 + \frac{\gamma \nu}{\sigma_0^2}$, which matters for the non-nesting peaks. These points have already been noted by MS95. However, the meaning of the scale-functions of the nesting peaks have not been pointed out until this paper with the help of the graphical picture in the skeleton tree formalism.
Figure 3. The scale-functions of various peaks are presented for various fluctuation models of the initial density fluctuation. The solid, dashed and dotted lines represent the scale-functions of the non-nesting peaks, all the peaks, and the nesting peaks, respectively.

4.4 Scale function of non-nesting peaks

The scale-function of the non-nesting peaks \( N(R, \delta_c) \) can be obtained by subtracting that of the nesting peaks \( N_{pk}^n(R, \delta_c) \) from that of all the peaks \( N_{pk}(R, \delta_c) \), which includes the nesting peaks \( N_{pk}^n(R, \delta_c) \) at \( \delta_c \) on the scale \( R \). Then, the notation of the corrected scale-function is

\[
N(R, \delta_c) = N_{pk}(R, \delta_c) - N_{pk}^n(R, \delta_c)
\]

(36)

\[
= N_{pk}(\nu, R) \sigma_{\nu}(R) R \, dR,
\]

(37)

where \( \theta = \langle \delta \rangle_{pk} - x_0 \). The mapping relation for the non-nesting peaks between \( \nu \) and \( R \):

\[
d\nu = \frac{\sigma_{\nu}(R)}{\sigma_0} R \, dR.
\]

(38)

Then, the two expressions of \( N(R, \delta_c) \) and \( N_{pk}^n(R, \delta_c) \) differ by a factor of \( \frac{\theta}{\nu} \).

4.5 Comparison of scale functions

Fig. 3 illustrates the scale-functions of (respectively) the non-nesting peaks, all the peaks, and the nesting peaks derived in the above subsections: \( N(R, \delta_c)R_0^3, N_{pk}(R, \delta_c)R_0^3 \), and \( N_{pk}^n(R, \delta_c)R_0^3 \) versus \( R/R_0 \), where we defined \( R_0 \) as the typical collapsing scale as determined by \( \delta_c / \sigma_0(R_0) = 1 \) with \( \delta_c = 1.69 \) and the standard cold dark matter (SCDM) model, use the normalization of 8 Mpc with the bias factor \( b = 1 \). Those of the non-nesting peaks, all the peaks, and the nesting peaks are represented by the solid, dashed and dotted lines, respectively. The number of nesting peaks is negligible at scales smaller than \( R_0 \), it becomes comparable to that of all the peaks and the non-nesting peaks at the larger scale. It shows a natural feature, i.e. that the nesting process becomes active with mergers on a scale larger than the typical collapsing scale \( R_0 \).

Fig. 3 shows that the scale-functions \( N_{pk}(R, \delta_c)R_0^3 \) become proportional to \( R^{-3} \) asymptotically in the small scale. As shown in BBKS, the cumulative number is a useful quantity which can be evaluated analytically,

\[
n_{pk}(\nu = -\infty) = \int_{\nu = -\infty}^{\infty} d\nu N_{pk} = \frac{29 - 6\sqrt{6}}{5\sqrt{2}/2\pi} R_0^3.
\]

(39)

As \( R \to 0 \), the density contrast \( \nu \to 0 \) and the peaks of the low contrast \( \nu \) dominate in the cumulative number. Then, \( N_{pk}(R, \delta_c)R_0^3 \) is \( \approx \eta_{pk}(\nu \to 0; R_0) \) in the small scale. Since \( R_0 \) is proportional to the resolution scale \( R \), the asymptotic feature seen in Fig. 3 is not unexpected. The two functions in the small mass are different from each other between the peak theories like the skeleton tree formalism and the PS formalism as pointed out by Appel & Jones (1990). Some differences also appear in the concept between them. These differentiations will be discussed, along with the topics of mass functions, in the next section.

5 THE MASS FUNCTION

We have not yet presented any mass functions, even if they are almost identical to the scale-functions shown above. In this section, we will derive the mass functions and compare them to the results from other models, like the PS formalism.

The mass function \( N(M, t) \) is often derived as the unconditional mass function \( f(\nu) \) which represents the fraction of the mass collapsed into the objects between \( M \) and \( M + dM \) at the collapsed threshold \( \delta_c(t) \) as

\[
\frac{M}{\rho_0} N(M, t) \, dM = f(\nu) \left| \frac{d\nu}{dM} \right| \, dM,
\]

(40)

where \( \rho_0 \) is the background density and \( \nu = \delta_c(t)/\sigma(M) \) is the ratio of the collapsed threshold to the rms fluctuation of the object \( M \). The function \( f(\nu) \) seems to be a universal form that is independent of redshift and power spectrum, even though we could not learn why this scaling holds for more general cases in numerical simulations (e.g. Sheth & Tormen 1999). We will also express the unconditional mass function in skeleton tree formalism, to compare it with those of other formalisms. In order to derive the unconditional mass function in the skeleton tree formalism, we start the first subsection with the mass estimation from the resolution scale.

5.1 Mass assignment with a gaussian filter

The physical basis of the peak theories is that the material around a peak that is smoothed with a Gaussian filter of a scale \( R_G \) is contained in a halo with mass \( \propto \rho_0 R_G^3 \), and this filter smoothing can also describe the accretion and the merging processes. This picture seems reasonable as discussed above; however, this Gaussian filtering shares the deficiency that an assigned mass \( M \) cannot be guaranteed to the appropriate volume \( V_{R_G} = 4\pi/3 R_G^3 \) in the scale that the top-hat filtering of a scale \( R_{th} \), with which the power spectrum amplitude in the models can naturally be compared with the observed or simulated ones. There is a contrast between a peak theory like the skeleton tree formalism with a Gaussian filter, and the PS formalism with a top-hat filter, when we define their mass functions.
While the PS formalism can automatically be normalized with an assumption for all the mass contained in the collapsed halo, there are a variety of choices for peak mass estimates. The simplest choice, for the volume with the Gaussian filter of the scale \( R_{G_0} \), is \( (2\pi)^{3/2} R_{G_0}^3 \). However, the collapse of this mass requires that the filtered density \( F(r; R_i) > \delta_i \) at only one point in the peak region. This condition seems artificial. While spherical top-hat filtering can account for a reasonable collapse condition, with a physical connection between the linear density growth and the spherical collapse solution \( \delta_c = 1.69 \), the peak with the Gaussian filtering can be higher than that averaged with the top-hat filter in the same collapsed region. Then, let consider the equivalent sphere with Gaussian filtering of a scale \( R_{G_0} \), which reproduces the same value of \( \sigma_0 \) with the top-hat filtering of a scale \( R_{th} \). This can introduce a proportional relation \( R_{th} = q_{SK} R_{G_0} \). The volume of the equivalent sphere is then represented with a modification parameter \( q_{SK} \) as

\[
\text{Vol}_{SK}(R_{G_0}) = \frac{4\pi}{3} (q_{SK} R_{G_0})^3. \tag{41}
\]

The exact relation between \( R_{th} \) and \( R_{G_0} \) for an equal \( \sigma_0 \) depends on the power spectrum as shown in Appendix D. For the power spectra of interest to us, however, the relation between \( R_{th} \) and \( R_{G_0} \) is approximated with \( q_{SK} = 2.3 \) at \( \delta_c = 1.69 \). It means that we can assign a uniformly larger mass, which is a factor of \( \approx 3 \) larger than \( (2\pi)^{3/2} R_{G_0}^3 \). Note that this mass assignment from the condition of equal \( \sigma_0 \) at a fixed threshold is basically identical to rescaling to a higher threshold for the mass assignment with \( R_{th} = R_{G_0} \).

The skeleton tree formalism will, then, obtain the mass assignment by adapting the rms fluctuation for the Gaussian filtering as being equal to the top-hat filtering. On the other hand, another mass assignment has been introduced with the normalization condition for mass functions,

\[
\int_0^\infty \mathrm{d}\nu \, f(\nu) = 1. \tag{42}
\]

This implies that the modification parameter \( q_{SK} \) is a constant in the case of the power-law spectrum, or as a near constant even in the case of a non-power-law spectrum like the CDM, as discussed in the CUSP formalism. The mass assignment arising from the normalization condition also depends on the collapsed threshold. It suggests that the mass assignment in the CUSP formalism from the normalization condition can be degenerate as compared with that from the equal rms fluctuation, as we mentioned above.

We also wish to remark upon the normalization condition, in which all the mass in the universe is to exist in bound objects of some mass, however small. This physical reasoning is still unclear when compared to the spherical collapsing picture introduced above.

In this paper, we used the equal \( \sigma_0 \) model of spherical collapse as our method of mass estimation in skeleton tree formalism.

### 5.2 Comparison of mass functions

We present the unconditional mass functions from the skeleton tree formalism and the other main formalisms in Fig. 4. Using the scale function of the non-nesting peaks, the unconditional mass function in skeleton tree formalism is given as

\[
\hat{s}_{sk}(\nu) = \text{Vol}_{sk}(R) V(R, \hat{\nu}) \frac{|R|}{|\partial R/\partial \nu|}. \tag{43}
\]

The PS formula has a universal shape as

\[
f_{PS}(\nu) \, d\nu = \frac{2A}{\sqrt{2\pi}} \exp \left[ -\frac{\nu^2}{2} \right] d\nu. \tag{44}
\]

For the case \( A = 1 \), the PS mass function can automatically satisfy the normalization condition. Although the emphasis hereof has been on how well the PS formalism reproduces the mass function in numerical simulations, this agreement is by no means perfect. The simulations have more massive haloes, and fewer intermediate and small-mass haloes, than are predicted by the PS formalism. As remarked by Sheth & Tormen (1999), in the simulations of clustering SCDM, OCDM and \( \Lambda \)CDM models by the GIF/Virgo Consortium collaboration (e.g. Kauffmann et al. 1999), the mass function has a universal shape called the GIF mass function,

\[
f_{GIF}(\nu) \, d\nu = 2A \left( 1 + \frac{1}{\nu^2} \right) \exp \left[ -\frac{\nu^2}{2} \right] \frac{1/2}{\nu^2} \, d\nu, \tag{45}
\]

where \( \nu' = \sqrt{\bar{a} \nu}, a = 0.707, q = 0.3 \) and \( A = 0.322 \).

We can see the similarities and differences between the mass functions of the skeleton tree formalism, the PS formalism, and the GIF mass function. While the skeleton tree formalism reproduces the GIF mass function in more massive haloes better than the PS formalism, there is much less agreement when considering small-mass haloes. The good agreement of the skeleton tree formalism with the GIF function at high contrast \( \nu > 1 \) suggests that peaks can indeed follow collapsed objects, as the statistics of the initial condition can be reserved during the quasi-linear stage of \( \nu > 1 \). The disagreement at low contrast \( \nu < 1 \) suggests that non-linear dynamics affects the statistics. This seems natural, since objects smaller than \( R_0 \), corresponding with \( \nu < 1 \), are in non-linear growth in the hierarchical clustering scenario with gravitational instability. In that case, the skeleton tree formalism should include spatial non-linear effects. In this paper, however, we will not
attempt to reproduce the mass function in the small-mass range more accurately, as the improvements required for solving this non-linear effect are still a future target for skeleton tree formalism.

6 EVOLUTION WITH DESTRUCTION AND REFORMATION IN MERGER AND ACCRETION

6.1 Scale growth with accretion and coalescence in mergers

We will consider the continuous scale growth rate with accretion for a peak of $R$. The scale growth rate with accretion can be described with the scaled Laplacian $x$ as

$$\hat{R}_{pk}(R, x, \delta_t) \, dt = \frac{1}{x \sigma R} \, d\delta_t. \quad (46)$$

We should note that this accretion growth rate for a peak depends on its particular value $x$. According to MS96, the mean growth rate for the objects of the scale $R$ can be expressed as

$$\hat{R}_{pk}(R, t) \, dt = \left\langle \frac{1}{x \sigma R} \right\rangle \, d\delta_t = \frac{1}{\langle x \rangle_{pk} \sigma R} \, d\delta_t, \quad (47)$$

where $\langle \rangle$ means the average of the function with $x$, and we used the relation

$$\langle x^{-1} \rangle = \frac{\int_0^\infty dx \, x^{-1} N_{pk}(R, \delta_t, x) \, dR}{\int_0^\infty dx \, N_{pk}(R, \delta_t, x) \, dR} = \frac{\int_0^\infty dx \, x^{-1} N_{pk}(x, R) \, dR}{\int_0^\infty dx \, N_{pk}(x, R) \, dR} = \langle x \rangle_{pk}^{-1}. \quad (48)$$

From this mean scale-growth rate, we can define the relative growth rate of the accretion as

$$r^a(R, t) \, dt = \frac{\partial \hat{R}_{pk}(R, t')}{\partial R}. \quad (49)$$

As we discussed in the previous section, nesting peaks generally appear in mergers. Furthermore, the nesting process induces mass growth through the coalescence of the smaller destroyed objects and the larger objects, which is quite different from continuous accretion growth. We will also consider the scale-growth rate averaged for all the peaks, taking into account the coalescence of the nesting peaks of $R$: $R_{ns}(R, \delta_t) \, dt$. The total volume growth of the coalescent peaks of $R$ can be represented by both sets of the number density for all the peaks of $R$ with the averaged growth rate $\hat{R}_{ns}$ and that for the coalescent peaks with the averaged accretion rate $\hat{R}_{pk}$ as

$$N_{pk}(R, \delta_t) \, 4\pi R^2 \hat{R}_{ns}(R, \delta_t) \, dt = N_{pk}^{\text{av}}(R, \delta_t) \, 4\pi R^2 \hat{R}_{ns}(R, \delta_t) \, dt,$$

where we consider that the averaged accretion includes the coalescence. From this relation, we obtain the growth rate of the coalescence,

$$\hat{R}_{ns}(R, \delta_t) \, dt = \frac{N_{pk}^{\text{av}}(R, \delta_t)}{N_{pk}(R, \delta_t)} \hat{R}_{pk}(R, \delta_t) \, dt = \frac{\gamma \nu}{\langle x \rangle_{pk}} \hat{R}_{pk}(R, \delta_t) \, dt. \quad (51)$$

With the growth rate of the coalescence, we can also introduce the growth rate of pure accretion as

$$\hat{R}(R, \delta_t) \, dt = \hat{R}_{pk}(R, \delta_t) \, dt - \hat{R}_{ns}(R, \delta_t) \, dt. \quad (52)$$

We note that the growth rates of the coalescence and the pure accretion do not exceed that of the accretion originally introduced, since $\langle x \rangle_{pk} = \gamma \nu + \theta$ with $\theta > 0$.

These growth rates can be seen in Fig. 5. We can see that the coalescence (accretion) dominates on the high- (low-) mass scale. This is consistent with the picture that the objects form, often with the nesting coalescence occurring in mergers at high-mass scale, which can be observed in hierarchical clustering simulations.

6.2 Conservation formulae for scale functions

The conservation equation for the scale-function of objects $N_s(R, t)$ is in general given as

$$\frac{\partial N_s(R, t)}{\partial t} + \frac{\partial \hat{R}_s(R, t) N_s(R, t)}{\partial R} = S_s(R, t), \quad (53)$$

where the suffix in $x$ can be specified for either as all the peaks or as non-nesting peaks, $\hat{R}_s(R, t)$ is the scale-growth rate with accretion, and $S_s(R, t)$ can be given as the net source term with regard to destruction and reformation. This net source term can be represented with the instantaneous scale-functions of the disappearing and reforming peaks in the volume of $dR\delta_t$ as

$$S_s(R, t) \, dR \, d\delta_t = N_s^d(R, \delta_t) \, dR \, d\delta_t - N_s^r(R, \delta_t) \, dR \, d\delta_t, \quad (54)$$

where the superscripts represent reformation and destruction. Then, if we obtain the instantaneous scale-functions of destruction and reformation, and the scale-growth rate for the objects of interest, we can reproduce the evolution of the objects.

First, we will describe the conservation equation of all the peaks. As shown in Fig. 2 (b), in the evolution of all the peaks, we do not need to count the nesting peaks. The conservation equation without
The coalescence picture of the nesting peaks is thus the case of pure accretion growth rate, we can rewrite this to describe the instantaneous scale-function of the reforming peaks as

\[
\frac{\partial N_{\text{pk}}(R,t)}{\partial t} + \frac{\partial [R_{\text{pk}}(R,t)N_{\text{pk}}(R,t)]}{\partial R} = N^i(R,\delta_i) |\delta_i|.
\]

In this case, with the pure accretion growth rate, we can rewrite this to describe the instantaneous scale-function of the reforming peaks as

\[
N^i(R,\delta_i) dR d\delta_i = N^d(R,\delta_i) dR d\delta_i + \left[ \frac{\partial N(R,t) dR}{\partial t} + \frac{\partial (R(t)N(R,t))}{\partial R} \right] |\delta_i|^{-1}.
\]

### 6.3 Instantaneous rates for accretion, destruction and reformation

The instantaneous destruction and reformation rates can be defined directly from their instantaneous scale-functions as

\[
r^i_{\text{st}}(R,t) = \frac{N^d(R,\delta_i) dR}{N_i(R,t)} \frac{d\delta_i}{dR},
\]

\[
r'^i_{\text{st}}(R,t) = \frac{N^i(R,\delta_i) dR}{N_i(R,t)} \frac{d\delta_i}{dR}.
\]

With the rates of \(r^i_{\text{st}}(R,t)\) and \(r'^i_{\text{st}}(R,t)\), we can represent the source term as

\[
S_i(R,t) = [r^i_{\text{st}}(R,t) - r'^i_{\text{st}}(R,t)]N_i(R,t).
\]
The evolution of growth rates at fixed masses for the SCDM model in an Einstein–de Sitter universe. The solid and dashed lines present the merger and the accretion growth rate \( R_a/R \times t \) and \( R_{ac}/R \times t \). In the figure, the lowest solid (highest dashed) curve at the left (right) is for \( M/M_0 = 10^{-3} \), and successive curves are for the relative mass ratio to the typical collapsing mass \( M_0 \) of \( M/M_0 = 10^{-2.5}, 10^{-2}, 10^{-1.5}, 10^{-1}, 10^{-0.5}, 1, \) and \( 10^{0.5} \).

Our formalism with the equality between the left-hand side and the right-hand side of the conservation equation.

Thus, the consistency of the skeleton tree formalism encourages us to proceed to describe the evolution of the non-nesting peaks. With the improved description from the conservation formula, in Fig. 7, the destruction rate \( r^d \partial t/\partial \delta_c \), the reformation rate \( r^f \partial t/\partial \delta_c \), and the shift rate \( r^s \partial t/\partial \delta_c \) are presented with the solid, short dash-dotted, short dashed, and long dashed lines, respectively.

The destruction rate reaches a maximum around the typical collapsing scale \( R = R_0 \) and rapidly decreases in the larger-scale range \( R > R_0 \), in which the rate of reformation dominates. It means that the merger makes objects of \( R > R_0 \) grow by means of the coalescence of objects of \( R < R_0 \), where the critical scale of \( R_0 \) is related to the present threshold \( \delta_c^* \). These properties of the dominating merger at larger than the critical scale \( R_0 \) have already been suggested in previous \( N \)-body calculations (e.g. Navarro, Frenk & White 1997).

From the viewpoint of the mass growth history for an object of the scale \( R \), the results in Fig. 7 should indicate that the dominant growth process switches from merger coalescence to accretion around the threshold of \( \delta_c = \sigma_0(R) \). We can see this feature directly in Fig. 8, which represents the evolution of the growth rate, with the reformation and the accretion at fixed scales. From these results, in general, the growth of a halo begins with the merger coalescence process, then switches to the accretion-dominated process. This feature also was noted from the simulations.

We should remark that the results in Fig. 8 are not the mean growth histories for the individual objects since the fixed scales cannot be directly related to the final scales of the objects at present, as a result of the successive destruction and reformation in conjunction with accretion. The mean growth history for an individual object will be described in following papers.

Figure 8. The evolution of growth rates at fixed masses for the SCDM model in an Einstein–de Sitter universe. The solid and dashed lines present the merger and the accretion growth rate \( R_a/R \times t \) and \( R_{ac}/R \times t \). In the figure, the lowest solid (highest dashed) curve at the left (right) is for \( M/M_0 = 10^{-3} \), and successive curves are for the relative mass ratio to the typical collapsing mass \( M_0 \) of \( M/M_0 = 10^{-2.5}, 10^{-2}, 10^{-1.5}, 10^{-1}, 10^{-0.5}, 1, \) and \( 10^{0.5} \).

7 CONCLUSION

We have derived an analytical expression for the statistics and the evolution of cosmic bound objects, using destruction and reformation in mergers, and accretion, as the point and line processes of the sloping saddles and peaks in the landscape which produces the tree structure. We have named our scheme the skeleton tree formalism. The skeleton tree formalism is derived as a natural improvement of the peaks theory of BBKS. In the landscape reproduced from the random field with a Gaussian filter, we have investigated the smoothing of the peaks as the accretion growth of objects, picked up the ‘sloping saddles’ as the destroyed objects in a merger, and the nesting peaks as pairs of a progenitor and a reformer in a merger.

This landscape concept of the Gaussian random field with Gaussian filtering in the skeleton tree formalism is similar to that in the CUSP formalism. However, we can improve it with a simple topological consideration of the critical points. We would like to note that the skeleton tree formalism can make clear the meaning of the nested process by distinguishing between all the peaks, the nesting peaks, and the non-nesting peaks, as these scale-functions of \( N_{pk}(R, \delta_0) \) \( dR \), \( N^{ns}_{pk}(R, \delta_0) \) \( dR \), and \( N(R, \delta_0) \) \( dR \) can be only derived from the number density of peaks of BBKS \( N_{pk}(\nu, R) \) \( d\nu \) with accurate mapping from a density contrast \( \nu \) to a scale \( R \). This distinction is also important when we try to obtain some quantities related to the spatial distribution of peaks, as Mo, Jing & White (1997) obtained the bias relation. Unfortunately, they used \( N_{pk}(\nu, R) \) \( d\nu \) without an accurate density of \( N_{pk} \), which differs by a factor of \( 1 + \frac{\Delta \delta}{\sigma} \); this is important for peaks with a smaller scale.

The skeleton tree formalism can calculate not only the scale-functions of these peaks with simple analytical forms, but also the instantaneous scale-function of the destroyed peaks by counting the sloping saddles. By contrast, in order to obtain this information, the CUSP formalism requires some stages of calculation as the first stage of obtaining the joint scale-function of the peaks nesting into a background peak, the second stage of the derivative with the density contrast in order to give the instantaneous joint scale-function of the merging peaks into a background peak, and the final stage of the integration with the background scale. Thus, the skeleton tree formalism is more convenient in some applications compared with the CUSP formalism since the calculation steps are reduced even if we calculate the correlations of the higher derivatives of the density field.

We have explained the meaning of the mass assignment with Gaussian filtering, which has not been clear until now. We have also compared the mass function of the skeleton tree formalism to others, like the Press–Schechter formalism and the GIF formulas that use the top-hat filter.

Our obtained rates of destruction and accretion for all the peaks in any hierarchical clustering models, can reproduce the evolution of the scale-function with the conservation accretion. This reproduction of the evolving scale-function verifies the self-consistency of the skeleton tree formalism. This self-calibration is another advantage for a theoretical formalism.

The skeleton tree formalism finds that the merging processes are efficient around the critical scale of \( R_0 \) determined as \( \delta_c^* = \sigma_0(R_0) \). The dominant growth process of the objects switches from merger coalescence to accretion around the critical collapsing scale. It is important to reproduce the mass-growth history and to make a distinction between merger coalescence and accretion when we attempt to reproduce cosmic structure and galaxy formation in a hierarchical scenario.
ACKNOWLEDGMENTS

The author is grateful to H.J. Mo and S.D.M. White for warm hospitality and useful discussions when HH visited MPA in Garching, and appreciates S. Nishizaki’s careful reading of the manuscript. This work was partially supported by a Grant-in-Aid from the Japanese Ministry of Education, Science and Culture, and the Sumitomo Foundation.

REFERENCES

Bond J. R., 1989, in Astbury A., Campbell B. A., Israel W., Khana F. C.,
eds, Frontiers in Physics: From Colliders to Cosmology. World
Scientific Press, Singapore, p. 182
Cole S., Aragón-Salamanca A., Frenck C. S., Navarro J. F., Zepf S., 1994,
MNRAS, 271, 781
Doroshkevich A. G., 1970, Astrofizika, 6, 30
Kauffmann G., Colberg J. M., Antonaldo D. W., White S. D. M., 1999,
MNRAS, 307, 529

APPENDIX A: THE COVARIANCE MATRIX

We introduce 20-dimensional vector \( F^{(20)} = (F, \nabla F, \nabla \otimes \nabla F, \nabla \otimes \nabla \otimes \nabla F) \). The \( \nabla F, \nabla \otimes \nabla F, \) and \( \nabla \otimes \nabla \otimes \nabla F \) have three, six, and ten independent components, respectively. They can be expressed in terms of Fourier transform as

\[
F(r, R) = \int d^3 k e^{i k \cdot r} F(k; R),
\]

(\ref{eq:F(r,R)})

\[
F_{a}(r, R) = \frac{\delta F(r, R)}{\delta x_a} = i \int d^3 k e^{i k \cdot r} k_a F(k; R),
\]

(\ref{eq:F_a(r,R)})

\[
F_{a \beta}(r, R) = \frac{\delta^2 F(r, R)}{\delta x_a \delta x_\beta} = -i \int d^3 k e^{i k \cdot r} k_a k_\beta F(k; R),
\]

(\ref{eq:F_a_b(r,R)})

\[
F_{a \beta \gamma}(r, R) = \frac{\delta^3 F(r, R)}{\delta x_a \delta x_\beta \delta x_\gamma} = -i \int d^3 k e^{i k \cdot r} k_a k_\beta k_\gamma F(k; R).
\]

(\ref{eq:F_a_b_gamma(r,R)})

It is useful to introduce the integrals over the filtered fluctuation spectrum

\[
\sigma^2_f(R) = 4 \pi \int_0^\infty dk k^{3/2} P(k; R),
\]

(\ref{eq:sigma_f(R)})

when transforming the above values to the non-dimensional ones.

The covariance matrix \( C^{(20 \times 20)} \) is defined as the expectation value of the direct product of the vector \( F^{(20)} \):

\[
C^{(20 \times 20)} = \langle F^{(20)} F^{(20)} \rangle.
\]

(\ref{eq:C(20x20)})

We can represent the matrix as

\[
C^{(20 \times 20)} = \begin{pmatrix}
\sigma_0^2 & 0 & M_{02}^T & 0 \\
0 & M_{11} & 0 & M_{13} \\
M_{02} & 0 & M_{22} & 0 \\
0 & M_{13} & 0 & M_{33}
\end{pmatrix},
\]

(\ref{eq:C_matrix})

\[
M_{11} = \frac{\sigma_0^2}{3} I^{(3 \times 3)},
\]

(\ref{eq:M_11})

\[
M_{22} = \frac{\sigma_0^2}{5} M_{22},
\]

(\ref{eq:M_22})

\[
M_{33} = \frac{\sigma_0^2}{7} \begin{pmatrix}
M_{33} & 0 & 0 & 0 \\
0 & M_{33} & 0 & 0 \\
0 & 0 & M_{33} & 0 \\
0 & 0 & 0 & 1/15
\end{pmatrix},
\]

(\ref{eq:M_33})

\[
M_{02} = -\frac{\sigma_0^2}{3} (K^{(3 \times 3)}, 0^{(3 \times 3)}),
\]

(\ref{eq:M_02})

\[
M_{13} = -\frac{\sigma_0^2}{5} (\tilde{M}_{13}, \tilde{M}_{13}, \tilde{M}_{13}, 0^{(3 \times 1)}),
\]

(\ref{eq:M_13})

\[
\tilde{M}_{13} = e_i \otimes (1, 1/3, 1/3), e_i = (\delta_{1i}, \delta_{2i}, \delta_{3i}),
\]

(\ref{eq:tilde_M_13})

where \( K^{(n \times m)}, I^{(n \times n)} \) and \( 0^{(n \times m)} \) are the \( (n \times m) \) dimensional matrix with every entry unity, the \( (n \times m) \) dimensional unit matrix and the \( (n \times m) \) dimensional null matrix, respectively.

APPENDIX B: CALCULATIONS OF JOINT PROBABILITIES

B1 Theorem for the joint probability expansion

According to Adler (1981) and BBKS, all Gaussian-distributed parameters represented by the \( n \)-dimensional vector \( Z = (Y, X) \), which can be divided into the \( m \)-dimensional vector \( Y \) and \( (n-m) \) dimensional vector \( X \), imply that the divided conditional probability \( p(Y|X) \) is a Gaussian

\[
p(Y|X) = \frac{\exp \left[ -\frac{1}{2} \left( \Delta Y \otimes \Delta Y \Delta X \right)^{-1} \Delta Y \right]}{(2\pi)^m/2 \left| \det (\Delta Y \otimes \Delta Y \Delta X) \right|^{1/2}},
\]

(\ref{eq:B1})
where \( \Delta Y = \Delta Y - \langle \Delta Y | \Delta X \rangle \), with the mean
\[
\langle \Delta Y | \Delta X \rangle = \langle \Delta Y \otimes \Delta X \rangle (\Delta X^2)^{-1} \Delta X, \tag{B2}
\]
and the conditional covariance matrix
\[
C(Y|X) = \langle \Delta Y \otimes \Delta Y | \Delta X \rangle \]
\[= \langle \Delta Y \otimes \Delta Y \rangle - \langle \Delta Y \otimes \Delta X \rangle (\Delta X^2)^{-1} \langle \Delta X \otimes \Delta Y \rangle, \tag{B4}
\]
where \( \langle \cdot \rangle \) represents the mean value and \( \Delta X = X - \langle X \rangle \). This is a general theorem which is extremely useful when we calculate the divided conditional probabilities.

### B2 Divided conditional probability

We can expand the joint probability distribution with the divided conditional probabilities as
\[
p(F, F_a, F_{a \beta}, F_{a \beta \gamma}) \, dF \, dF^{(3)} \, dF^{(6)} \, dF^{(10)} = p(F_{a \beta}|F_a, F_{a \beta}) \, dF^{(6)} \, dF^{(10)} \]
\[\times p(F, F_a, F_{a \beta}) \, dF \, dF^{(3)} \, dF^{(6)} \tag{B5}
\]
Using the theorem for the divided conditional probability, these divided conditional probability functions are represented in the explicit forms
\[
p(F, F_a) = \frac{3^{1/2}}{2\pi^{3/2}} \exp\left[-\frac{\epsilon}{3} Q(F, F_a)\right], \tag{B7}
\]
\[
p(F_{a \beta}|F_a, F_{a \beta}) = p(F_{a \beta}|F_a) = \exp\left[-\frac{\epsilon}{3} Q(F_{a \beta}|F_a)\right], \tag{B8}
\]
\[
p(F_{a \beta \gamma}|F_a, F_{a \beta}) = p(F_{a \beta \gamma}|F_a, F_{a \beta}) = \exp\left[-\frac{\epsilon}{3} Q(F_{a \beta \gamma}|F_a)\right], \tag{B9}
\]
where
\[
Q(F, F_a) = \frac{F^2}{2 \sigma^2} + \frac{3F_a F}{\sigma^2}, \tag{B10}
\]
\[
Q(F_{a \beta}|F_a) = F_{a \beta}^T C(F_{a \beta}|F_a)^{-1} F_{a \beta}, \tag{B11}
\]
\[
Q(F_{a \beta \gamma}|F_a) = F_{a \beta \gamma}^T C(F_{a \beta \gamma}|F_a)^{-1} F_{a \beta \gamma}, \tag{B12}
\]
\[
\hat{F}_{a \beta} = F_{a \beta} - M_{22} \sigma_0^2 F, \tag{B13}
\]
\[
\hat{F}_{a \beta \gamma} = F_{a \beta \gamma} - 3 \sigma_1^2 M_{13} F, \tag{B14}
\]
and the conditional covariance matrix
\[
C(F_{a \beta}|F_a) = M_{22} - M_{22} \sigma_0^2 M_{22} = \sigma_1^2 (M_{22} - M_{202}), \tag{B15}
\]
\[
\hat{M}_{202} = \frac{5 \sigma_1^2}{9 \sigma_2^2 \sigma_0^6} \left( K^{(3 \times 3)} \ 0^{(3 \times 3)} \ 0^{(3 \times 3)} \right), \tag{B16}
\]
\[
C(F_{a \beta \gamma}|F_a) = M_{33} - M_{33} \sigma_1^2 \sigma_1^2 M_{13}. \tag{B17}
\]

### B3 Probability function for peaks

For the peaks, the original covariance matrix for the probability is represented in the 10-dimensional vector \( \mathbf{F}^{(10)} \). As presented by BBKS, we can see that the three degrees of freedom related to the direction drop off from the parameter space for homogeneous fields. After introducing the eigen coordinate, we can represent the peak number density in 7-dimensional parameter space.

For convenience, we will introduce a new set of variables,
\[
\sigma_{0 \nu} = F, \tag{B18}
\]
\[
\sigma_{1 \nu} = F_a, \tag{B19}
\]
\[
\sigma_{2 x} = -\nabla^2 F = \lambda_1 + \lambda_2 + \lambda_3, \tag{B20}
\]
\[
\sigma_{2 y} = \frac{(\lambda_1 - \lambda_3)}{2}, \tag{B21}
\]
\[
\sigma_{2 z} = \frac{(\lambda_1 - 2 \lambda_2 + \lambda_3)}{2}, \tag{B22}
\]
\[
dF_{a \beta}^6 = |(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3)| d\lambda_1 d\lambda_2 d\lambda_3 \times d \text{vol}[SO(3)]
\]
\[= |2 y(y^2 - z^2)| \frac{2}{3} \sigma_1^2 \, dx \, dy \, dz \, d \text{vol}[SO(3)], \tag{B23}
\]
where \( d\text{vol}[SO(3)] \) is the volume element of the three-dimensional rotational group \( SO(3) \). We use the relations
\[
\lambda_1 = \frac{\sigma_2}{3} (x + 3y + z), \tag{B24}
\]
\[
\lambda_2 = \frac{\sigma_2}{3} (x - 2z), \tag{B25}
\]
\[
\lambda_3 = \frac{\sigma_2}{3} (x - 3y + z), \tag{B26}
\]
\[
\sum_{i=1}^3 d\lambda_i = \frac{2}{3} \sigma_1^2 \, dx \, dy \, dz. \tag{B27}
\]

With these variables of \((\nu, \eta, x, y, z)\) with \( \eta = (\nu_1, \nu_2, \nu_3) \) from the original variables of \((F, F_a, F_{a \beta})\), the probability function in these is represented as
\[
p (F, F_a, F_{a \beta}) \, dF \, dF_a \, dF_{a \beta} = p(\nu, \eta, x, y, z, \alpha, \beta, \gamma) \, d\nu \, d\eta \tag{B28}
\]
\[\times |2 y(y^2 - z^2)| \frac{2}{3} \sigma_1^2 \, dx \, dy \, dz \, d \text{vol}[SO(3)],
\]
where \( \alpha, \beta, \gamma \) are the Euler’s angles. The directional dependence is not so important in the isotropic field. The mean of the peak number density is sufficient for us to consider the statistics of the density field. The mean can be obtained by the angle independent
of integrations.

\[ \int \text{dvol}[SO(3)] p(\nu, \eta, x, y, z, \phi, \psi) = \frac{2\pi^2}{3} p(x, y, z) p(\nu, \eta), \]

where we used \( \int \text{dvol}[SO(3)] = 2\pi^2/3! \) with the degeneracy of the triad orientation of the axis for the eigenvalues.

The determinant and the inverse of the conditional covariance matrix are given explicitly as

\[ \text{ldet} \mathbf{C}(\nu_{34})^{1/2} = \frac{2}{5^{3/2} \times 3^3} (1 - \gamma^2)^{1/2}, \quad \gamma = \frac{\sigma^2_{12}}{\sigma_{12} \sigma_{13}}, \]

where

\[ \mathbf{C}(\nu_{34})^{-1} = \begin{pmatrix} \frac{1}{\gamma^2} \mathbf{K}^{(3\times3)} + \mathbf{R} & \mathbf{0} \\ \mathbf{0} & -15 \mathbf{I}^{(3\times3)} \end{pmatrix}, \]

\[ \mathbf{R} = 5 \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix}. \]

The probability functions \( p(x, y, z|\nu) \), \( p(\nu) \) are the same as those described by BBKS as

\[ p(\nu, \eta) = \frac{3^{3/2}}{(2\pi)^2} \exp \left[ -\frac{1}{2} (\nu^2 + 3 \eta \times \eta) \right], \]

\[ p(x, y, z|\nu) = \frac{3 \times 5^{2/3}}{(2\pi)^3 \times 2 \times (1 - \gamma^2)^{1/2}} \exp \left[ -\frac{1}{2} Q(x, y, z) \right], \]

\[ Q(x, y, z) = \frac{(x - x_0)^2}{(1 - \gamma^2)} + 15y^2 + 5z^2, \quad x_0 = \gamma \nu. \]

### B4 Probability function for sloping saddles

For the sloping saddles, in general, we should start from the original covariance matrix in the 20-dimensional vector with 10 variables of the third derivatives of the field added to that for the peaks. After introducing the eigen coordinate and rotating to this coordinate, however, we need the third derivatives only along the degenerate direction as the additional variables \( \nu_{34} = F_{34\alpha}/\alpha_3 \). Similarly, as shown in the case of the peak, the statistics of the sloping saddles of the homogeneous field can finally be determined in the parameter space \( (\nu, \eta, x, y, z, w) \), where \( \mathbf{w} = (w_1, w_2, w_3) = (\nu_{311}, \nu_{322}, \nu_{333}) \). The divided conditional probability with the third derivatives of \( \mathbf{w} \) can be represented as

\[ p(\mathbf{w}|\nu_{34}) = \frac{\exp \left[ -\frac{1}{2} Q(\mathbf{w}|\nu_{34}) \right]}{(2\pi)^{10} \text{ldet} \mathbf{C}(\nu_{34})^{1/2}}, \]

\[ Q(\mathbf{w}|\nu_{34}) = \mathbf{w}^T \mathbf{C}(\nu_{34})^{-1} \mathbf{w}, \]

\[ \mathbf{w} = -\frac{3}{5} \sqrt{\nu_3} \mathbf{M}^T \]

\[ \kappa = \frac{\sigma^2_{12}}{\sigma_{13} \sigma_{14}}. \]

The determinant and the inverse of the conditional covariance matrix for the variables \( \nu_{34} \) are represented as

\[ \text{ldet} \mathbf{C}(\mathbf{w}|\nu_{34})^{1/2} = \frac{2}{5^{1/2} \times 3 \times 7} (1 - \kappa^2)^{1/2}, \]

\[ \mathbf{C}(\mathbf{w}|\nu_{34})^{-1} = \frac{1}{(1 - \kappa^2)^2} \begin{pmatrix} c_3 & c_o & c_o \\ c_o & c_2 & c_o \\ c_o & c_o & c_1 \end{pmatrix}, \]

\[ c_3 = 10 - 7\kappa^2, \]

\[ c_2 = c_1 = 45 - 42\kappa^2, \]

\[ c_o = \frac{21 \kappa^2 - 15}{2}. \]

With the condition of the first derivatives \( F_3 \) for the sloping saddles, we can represent the exponents parts of the divided conditional probability as

\[ Q(\mathbf{w}|\nu_{34}) = Q_1(w_1|w_2, w_3, \nu_3) + Q_2(w_2|w_3, \nu_3) + Q_3(w_3|\nu_3), \]

\[ Q_1(w_1|w_2, w_3, \nu_3) = \frac{3(15 - 14\kappa^2)}{(1 - \kappa^2)} \]

\[ \times \left[ w_1 - \frac{(5 - 7\kappa^2)}{2(15 - 14\kappa^2)} (w_3 + \bar{w}_2) \right]^2, \]

\[ Q_2(w_2|w_3, \nu_3) = \frac{3 \times 5 \times 7(25 - 21\kappa^2)}{2(15 - 14\kappa^2)} \]

\[ \times \left[ w_2 - \frac{(5 - 7\kappa^2)}{2(25 - 21\kappa^2)} w_3 \right]^2, \]

\[ Q_3(w_3|\nu_3) = \frac{5^2 \times 7}{(25 - 21\kappa^2)^2} w_{33}. \]

We will list some integrations as

\[ I_1 = \int_{-\infty}^{\infty} \text{dw}_1 \exp \left[ -\frac{1}{2} Q_1(w_1|w_2, w_3, \nu_3) \right] = \frac{\sqrt{2\pi}(1 - \kappa^2)^{1/2}}{3^{1/2}(15 - 14\kappa^2)^{1/2}}, \]

\[ I_2 = \int_{-\infty}^{\infty} \text{dw}_2 \exp \left[ -\frac{1}{2} Q_2(w_2|w_3, \nu_3) \right] = \frac{2\sqrt{2\pi}(15 - 14\kappa^2)^{1/2}}{(3 \times 5 \times 7)^{1/2}(25 - 21\kappa^2)^{1/2}}, \]

\[ I_{22} = \int_{-\infty}^{\infty} \text{dw}_2 \text{dw}_3 \exp \left[ -\frac{1}{2} Q_2(w_2|w_3, \nu_3) \right] = \int_{-\infty}^{\infty} \text{dw}_2 \left[ \frac{(5 - 7\kappa^2)}{2(25 - 21\kappa^2)} w_3 \right] \exp \left[ -\frac{1}{2} Q_2(w_2|w_3, \nu_3) \right] = \frac{(5 - 7\kappa^2)}{(25 - 21\kappa^2)} w_3^2 I_2. \]
As shown in the definition of the sloping saddles, we are interested mainly in the probability for $w_3$ in the third derivative components. The divided conditional probability function of $w_3$ can be obtained from the integration with the other third derivative components $w_1, w_2$, as

$$p(w_3|v_3) = \int_{-\infty}^{\infty} dw_1 \int_{-\infty}^{\infty} dw_2 p(w_1, v_3) dw_3$$

$$= \frac{5^{3/2} \times 3 \times 7}{2(2\pi)^{1/2} (1 - \kappa^2)^{3/2}} \times I_1 \times I_2$$

$$\times \exp \left[ -\frac{1}{2} \tilde{Q}(w_3) \right] dw_3$$

$$= \frac{5 \times 7^{1/2}}{\sqrt{2\pi(25 - 21\kappa^2)}} \times \exp \left[ -\frac{1}{2} \tilde{Q}(w_3) \right] dw_3, \quad (B52)$$

$$\tilde{Q}(w_3) = \frac{5^2 \times 7}{(25 - 21\kappa^2)} \left( w_3 + \frac{3}{5} \kappa p_3 \right)^2. \quad (B53)$$

We can see a reasonable normalization in the integration of $p(w_3|v_3)$ with $w_3$ as

$$\int_{-\infty}^{\infty} dw_3 p(w_3|v_3) = 1. \quad (B54)$$

For convenience, we will represent the probability of $w_3$ with $v_3 = 0$,

$$p(w_3|v_3 = 0) = \frac{1}{\sqrt{2\pi \sigma_{w_3}}} \exp \left[ -\frac{1}{2} \frac{w_3^2}{\sigma_{w_3}^2} \right], \quad (B55)$$

$$\sigma_{w_3}^2 = \frac{(25 - 21\kappa^2)}{5^2 \times 7}. \quad (B56)$$

We calculate the integration required in the calculation of the instantaneous scale-function of sloping saddles,

$$\int d^3 w_3 (w_1 + w_2 + w_3) p(w_1, v_3) = (w_3^2)_a (1 + 2q_{ss}). \quad (B57)$$

$$(w_3^2)_a = \int_{-\infty}^{\infty} dw_3 (w_3) p(w_3|v_3) = \frac{(25 - 21\kappa^2)}{5^2 \times 7}, \quad (B58)$$

$$q_{ss} = \frac{(5 - 7\kappa^2)}{(25 - 21\kappa^2)}, \quad (B59)$$

where we used

$$\int_{-\infty}^{\infty} dw_1 w_1 \int_{-\infty}^{\infty} dw_2 p(w_1, v_3) dw_3 = \int_{-\infty}^{\infty} dw_2 p(w_1|v_3) dw_3$$

$$= q_{ss} \tilde{w}_3 p(w_3|v_3) dw_3. \quad (B60)$$

As discussed in Section 2, the point processes of peaks and sloping saddles can be defined in the 3D spatial coordinate and the 4D space extended with $R$, respectively. Then, while the ensemble-averaged density of peaks is the differential density per infinitesimal ranges of the resolution scale $R$ at a fixed density threshold $\delta$, the instantaneous ensemble-averaged density of the sloping saddles is the differential density per infinitesimal volume of $dR \cdot d\delta$. In the derivation of ensemble-averaged densities, we should note their differences from density counting spaces.

**APPENDIX C: CALCULATIONS OF ENSEMBLE-AVERAGED DENSITIES OF CRITICAL POINTS**

**C1 Constraints of peaks and sloping saddles**

Under the transformation to the diagonal principal axis, the peaks require the condition of $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$, which should be added to the above condition of the extrema (equation 14). With the additional condition, the constraint for the spatial distributions of the peaks is

$$C(F^{\text{peak}}) = \sigma_{w_3}^2 \lambda_1 \lambda_2 \lambda_3 \prod_{i=1}^{3} \delta(F_n) \prod_{i=1}^{3} \theta(\lambda_i), \quad (C1)$$

where $\theta(\lambda_i)$ is the Heaviside function.

The sloping saddles requires the additional condition of $\lambda_i > 0$ ($i = 1, 2$) as excluding the merger of the hole. With the constraint of equation (22), the constraint for the instantaneous spatial distribution of the sloping saddle can be described as

$$C(F^{\text{saddle}}) = \sigma_{w_3}^2 \sigma_x^2 \delta(F_n) \delta(F_n) \prod_{i=1}^{2} \delta(\lambda_i) \prod_{i=1}^{3} \theta(\lambda_i), \quad (C2)$$

where we use $\sigma_{w_3}^2 = F_{3ax}$.

**C2 Ensemble-averaged density of peaks**

The constraint of the peaks can be rewritten as

$$C(F^{\text{peak}}) = |\lambda_1 \lambda_2 \lambda_3| \prod_{i=1}^{3} \delta(F_n) \Theta(\lambda_i)$$

$$= (x - 2z)(x + z)^2 (3y)^2 \left( \frac{\sigma_2}{\sigma_1} \right)^3 \times \prod_{i=1}^{3} \delta(\nu_i) \Theta(x - 2z). \quad (C3)$$

Using this constraint, in order to derive the ensemble-averaged number density of peaks, we start from its probability-weighted number density,

$$n_{pk} (v, x, y, z; R) d\nu dx dy dz$$

$$= 6 \times p(v, \eta = 0, x, y, z) C(F^{\text{peak}}) d\nu dx dy dz$$

$$= \frac{5^{5/2} 3^{3/2}}{2(2\pi)^{3}} \left( \frac{\sigma_2}{\sigma_1} \right)^3 \frac{1}{\exp \left[ -\frac{1}{2} \nu^2 \right]} \exp \left[ -\frac{1}{2} Q(x, y, z) \right] \sqrt{1 - \gamma^2}$$

$$\times \phi_{pk}(x, y, z) \phi_{pk}(x, y, z) d\nu dx dy dz, \quad (C4)$$
\[ \psi_{pk}(x, y, z) = \frac{3^3}{2} \sigma_2^3 \lambda_1 \lambda_2 \lambda_3 (\lambda_1 - \lambda_2) (\lambda_2 - \lambda_3) (\lambda_1 - \lambda_3) \]
\[ = (x - 2z)(x + z)^2 - (3y)^2 (y^2 - z^2), \quad (C5) \]

where the condition is \( \lambda_i > 0 \) and we have multiplied the probability expression by 6 to account for the ordering of the eigenvalues of \( \lambda_i \).

After the integration over \( y,z \), the ensemble-averaged density of the peaks is given by

\[ \langle \eta_{pk}(\nu, x; R) \rangle d\nu dx = \frac{\exp \left[ -\frac{1}{2} \nu^2 \right]}{(2\pi)^2 R_\nu^2} \]
\[ \times f_{pk}(x) \exp \left[ -\frac{(x - x_0)^2}{2(1 - \gamma^2)} \right] \frac{dx}{\sqrt{2\pi(1 - \gamma^2)}}, \quad (C7) \]

where \( R_\nu = \sqrt{3}(\sigma_\nu / \sigma_2) \) and

\[ f_{pk}(x) = \frac{3^{5/2}}{2\sqrt{\pi}} \int_0^{\nu^2} dy e^{-15y^2/2} \int_0^{\nu y} \psi_{pk}(x, y, z) dz e^{-5z^2/2} \]
\[ + \int_0^{\nu^2} dy e^{-15y^2/2} \int_{x-y}^{\nu y} \psi_{pk}(x, y, z) dz e^{-5z^2/2} \]
\[ = \left( \frac{x^3 - 3x}{2} \right) \left\{ \text{Erf} \left[ (5/2)^{1/2} x \right] + \text{Erf} \left[ (5/2)^{1/2} \frac{x}{y} \right] \right\} \]
\[ + \left( \frac{2}{3\pi} \right)^{1/2} \left[ \frac{31x^2}{4} + \frac{8}{3} \right] e^{-5y^2/8} + \left( \frac{x^2}{2} - \frac{8}{3} \right) e^{-5y^2/2} \].

The density of the peaks of \( \nu \) at the scale \( R \) can be represented as

\[ \langle \eta_{pk}(\nu, x; R) \rangle d\nu dx = \frac{\exp \left[ -\frac{1}{2} \nu^2 \right]}{(2\pi)^2 R_\nu^2} G_{pk}(\gamma, x_0) d\nu, \quad (C9) \]

where

\[ G_{pk}(\gamma, x_0) = \int_0^{\infty} dx f_{pk}(x) \frac{\exp \left[ -\frac{(x - x_0)^2}{2(1 - \gamma^2)} \right]}{\sqrt{2\pi(1 - \gamma^2)}}. \quad (C10) \]

The function \( G_{pk}(\gamma, x_0) \) has the following fitting formula obtained by BBKS,

\[ G_{pk}(\gamma, x_0) = \frac{w^3 - 3y^2w + (Bw^2 + C_1) \exp[-Aw^2]}{1 + C_2 \exp[-C_3w]}, \]

\[ A = \frac{5/2}{(9 - 5\gamma^2)}, \quad (C12) \]
\[ B = \frac{432}{(10\pi)^{2/2}(9 - 5\gamma^2)^5/2}, \quad (C13) \]
\[ C_1 = 1.84 + 1.13(1 - \gamma^2)^{5/2}, \quad (C14) \]
\[ C_2 = 8.91 + 1.27 \exp(6.51\gamma^2), \quad (C15) \]

\[ C_3 = 2.58 \exp(1.05\gamma^2). \]

We also have the averaged \( x \),

\[ H_{pk}(\gamma, x_0) = \int_0^{\infty} dx f_{pk}(x) \frac{\exp \left[ -\frac{(x - x_0)^2}{2(1 - \gamma^2)} \right]}{\sqrt{2\pi(1 - \gamma^2)}} \]

\[ = \langle \chi_{pk} \rangle G_{pk}(\gamma, x_0). \quad (C17) \]

We can easily see that the mean value \( \langle \chi_{pk} \rangle \) is \( H_{pk}(\gamma, x_0)/G_{pk}(\gamma, x_0) \), which can be also represented as an approximated form as derived by BBKS:

\[ \langle \chi_{pk} \rangle = \gamma \nu + \theta(\gamma, \gamma \nu), \quad (C18) \]

\[ \theta(\gamma, \gamma \nu) = \frac{\theta_1 + \theta_2 \exp[-\gamma \nu(\gamma \nu - 2)]}{\theta_1 + 0.45 + (\gamma \nu - 2)^{1/2} + \gamma \nu^2}, \quad (C19) \]

\[ \theta_1 = 3(1 - \gamma^2), \theta_2 = (1.216 - 0.9\gamma^2). \quad (C20) \]

C3 Ensemble-averaged density of sloping saddles

The constraints of the sloping saddles are rewritten to

\[ C(F^{(20)}_{s, \text{saddles}}) dR = \frac{1}{3} \frac{\sigma_\nu \sigma_\gamma}{\sigma_2^2} (x + 3y + z)(x - 2z) w_3 \]
\[ \times (w_1 + w_2 + w_3) \]
\[ \times \delta(z - 3y + x) \delta(3)(\eta) \frac{\sigma_\gamma}{\sigma_2} R dR. \quad (C21) \]

With the constraint, the probability-weighted number density of the sloping saddles is represented as

\[ n_{sa} (x, y, z, w; R) d\nu dx dz d\nu dz \]
\[ = 6 \times p(w|v_3 = 0) d^3 w p(x, y, z|\nu) dx dz p(\nu, \eta = 0) d\nu \]
\[ \times C(F^{(20)}_{s, \text{saddles}}) dR \]
\[ = \frac{3^{5/2} \times 5^{5/2}}{(2\pi)^{2/2}} \frac{\sigma_\nu \sigma_\gamma}{\sigma_2^2} \psi_{sa}(x, y, z) \phi_{sa}(x, y, z) \]
\[ \times \frac{\exp \left[ -\frac{1}{2} \nu^2 \right]}{\sqrt{1 - \gamma^2}} \frac{\exp \left[ -\frac{1}{2} \gamma \nu(\gamma \nu - 2) \right]}{\sqrt{1 - \gamma^2}} \]
\[ \times 4w_3(w_1 + w_2 + w_3) p(w|v_3 = 0) d^3 w \times \frac{\sigma_\gamma}{\sigma_2} R dR, \quad (C22) \]

\[ \psi_{sa}(x, y, z) = \gamma (y^2 - z^2)(x + 3y + z) \times (x - 2z) \delta(z - 3y + x), \]

\[ \phi_{sa}(x, y, z) = \delta_{pk}(x, y, z), \quad (C23) \]

where the factor 6 is needed to account for the ordering of \( \lambda_i \) as the same as that of the peaks. With the integration over \( y,z \), we can obtain the density of the sloping saddles (similar to that of the peaks).
\[ \langle n_{ss}(\nu, x, w; R) \rangle \, d\nu \, dx \, d^3 w \, dR \]
\[ = \frac{\exp\left[ -\frac{1}{2} \nu^2 \right]}{(2\pi)^2 R_\nu^4} \frac{\exp[-(x - x_w)^2]}{(1 - \gamma^2)^{\frac{1}{2}} \sqrt{2\pi(1 - \gamma^2)}} \]
\[ \times \text{hw}(w_1 + w_2 + w_3)p(w)w_3 = 0 \, d^3 w \left( \frac{\sigma_{\nu}}{\sigma_1} \right) R \, dR, \quad \text{(C25)} \]

where we used the relation \( \sigma_2 \sigma_3 / \sigma_1^2 = \sqrt{5} \kappa^{-1} R_\nu^{-1} \), and
\[ f_{ss}(\kappa, x) = A_{ss}(\kappa) \int_0^{\nu} d\nu \mathcal{V}(\nu; x) \]
\[ = A_{ss}(\kappa) \exp\left( -\frac{5/4\kappa^2}{2} \right) \]
\[ \times \left[ \frac{4}{5} \left( 1 - e^{-15/8\kappa^2} \right) + \frac{1}{5} \kappa^2 \left( \frac{3\kappa^2}{2^3} - \frac{1}{10} \right) \right], \quad \text{(C26)} \]

where
\[ A_{ss}(\kappa) = \frac{3^{3/2} \times 5^{5/2}}{\sqrt{2\pi}}. \quad \text{(C27)} \]

\[ \mathcal{V}(\nu; x) = e^{-15/8}/2 \int_{3\kappa-x}^{\nu} \psi_{ss}(x, y, z)dz \, e^{-3\kappa/2} \]
\[ = 18y^2(x - 2y^2)(4y - x) \times \exp\left[ -\frac{15y^2}{2} - \frac{5(3y - x)^2}{2} \right] \]
\[ = 18y^2(x - 2y^2)(4y - x) \times \exp\left[ -\frac{60(y - x)^2}{2} - 5/4\kappa^2 \right]. \quad \text{(C28)} \]

The density averaged with \( w \) is given as
\[ \langle n_{ss}(\nu, x, R) \rangle \, d\nu \, dx \, dR \]
\[ = \frac{\exp\left[ -\frac{1}{2} \nu^2 \right]}{(2\pi)^2 R_\nu^4} \frac{\exp[-(x - x_w)^2]}{(1 - \gamma^2)^{\frac{1}{2}} \sqrt{2\pi(1 - \gamma^2)}} \]
\[ \times f_{ss}(x) \frac{\exp\left( -\frac{(x - x_w)^2}{2(1 - \gamma^2)} \right)}{\sqrt{2\pi(1 - \gamma^2)}} \, dx \]
\[ \times \left( w^2 \right)_{ss}(1 + 2q_{wss}) \frac{\sigma_{\nu}(R)}{\sigma_{\nu}(R_\nu^2)} R \, dR. \quad \text{(C29)} \]

With the integration over \( x \), the averaged density is
\[ \langle n_{ss}(\nu, R) \rangle \, d\nu \, dR \]
\[ = \frac{G_{ss}(\gamma, x_w) - \frac{1}{2} \nu^2}{(2\pi)^2 R_\nu^4} \, d\nu \]
\[ \times \left( w^2 \right)_{ss}(1 + 2q_{wss}) \frac{\sigma_{\nu}(R)}{\sigma_{\nu}(R_\nu^2)} R \, dR, \quad \text{(C30)} \]

where
\[ G_{ss}(\gamma, x_w) = \int_0^{\infty} f_{ss}(x) \frac{\exp\left( -\frac{(x - x_w)^2}{2(1 - \gamma^2)} \right)}{\sqrt{2\pi(1 - \gamma^2)}} \, dx, \quad \text{(C31)} \]

The exact result of the integration is obtained as an analytical form,
\[ G_{ss}(\gamma, x_w) = \frac{A_{ss}(\kappa)}{\sqrt{2\pi(1 - \gamma^2)}} \left[ 2(1 - \gamma^2)x_w \right] \frac{12x_w^2}{(9 - 5\gamma^2)^3} \left[ (9 - 5\gamma^2) \right] \]
\[ + \frac{(39 - 55\gamma^2)}{5} e^{\frac{\gamma_x}{2(1 - \gamma^2)}} + \frac{2\pi(1 - \gamma^2)}{5(9 - 5\gamma^2)^3} \]
\[ \times \left[ \left( 24x_w^2(27 - 35\gamma^2)(9 - 5\gamma^2)x_w \right) \right. \]
\[ + \left( 783 - 1230\gamma^2 + 575\gamma^4(9 - 5\gamma^2)^2 \right) \]
\[ \times \left\{ 1 + \text{Erf} \left[ x_w \sqrt{2(1 - \gamma^2)} \left( \frac{2}{9 - 5\gamma^2} \right) \right] \right\} \cdot e^{\frac{\gamma_x}{2(1 - \gamma^2)}}, \quad \text{(C32)} \]

The mean value of \( x \) at the sloping saddles is
\[ \langle x \rangle_{ss} = \frac{H_{ss}(\gamma, x_w)}{G_{ss}(\gamma, x_w)}, \quad \text{(C33)} \]

\[ H_{ss}(\gamma, x_w) = \int_0^{\infty} dx f_{ss}(x) \frac{\exp\left[ -\frac{(x - x_w)^2}{2(1 - \gamma^2)} \right]}{\sqrt{2\pi(1 - \gamma^2)}}, \quad \text{(C34)} \]

where we can also obtain the analytical result of the integration
\[ H_{ss}(\gamma, x_w) = \frac{A_{ss}(\kappa)}{\sqrt{2\pi(1 - \gamma^2)}} \left[ 4(1 - \gamma^2) \right] \frac{24x_w^2}{(9 - 5\gamma^2)^3} \left[ 9(1 - \gamma^2)^2 \right] \]
\[ + \frac{2(99 - 115\gamma^2)x_w^2}{25(9 - 5\gamma^2)^3} \left[ 9(1 - \gamma^2)^2 \right] \]
\[ \times e^{\frac{\gamma_x}{2(1 - \gamma^2)}} \]
\[ + \left[ 2403 - 4950\gamma^2 + 2675\gamma^4 \right] \]
\[ \times \left\{ 1 + \text{Erf} \left[ x_w \sqrt{2(1 - \gamma^2)} \left( \frac{2}{9 - 5\gamma^2} \right) \right] \right\} \]
\[ \times x_w e^{\frac{\gamma_x}{2(1 - \gamma^2)}}, \quad \text{(C35)} \]

\[ \langle x \rangle_{ss} e^{\frac{\gamma_x}{2(1 - \gamma^2)}} = \frac{2}{375} \left[ 2\pi(1 - \gamma^2) \right] \]
\[ \times \left\{ 1 + \text{Erf} \left[ x_w \sqrt{2(1 - \gamma^2)} \left( \frac{2}{9 - 5\gamma^2} \right) \right] \right\} x_w e^{\frac{\gamma_x}{2(1 - \gamma^2)}}. \]
When the power spectrum can be represented as a simple power law, as
\[ |d_k|^2 = A k^n, \]
we can obtain analytical forms for the jth order spectral moment,
\[ \sigma_j(R) = A^{1/2} \sqrt{2 \pi} R \left( \frac{n + 3}{2} + j \right)^{1/2}, \]

\[ \sigma_{j,\text{ch}}(R) = Q_j(n) \sigma_{j,\text{ch}}(R), \]

\[ Q_j(n) = \frac{3^2 \times 2^{n-1} \pi \Gamma(1 - n - 2j)}{\Gamma(5/2 - n/2 - j) \Gamma^2(1 - n/2 - j)} . \]

where the Gs and Th in the subscripts mean the Gaussian filter and the top-hat filter, respectively.

Thus, we can also obtain useful relations for the calculations, as
\[ \frac{\sigma_j(R)}{\sigma_j(R)} = \frac{(2i + n + 1)!!}{2^{i-j}(2j + n + 1)!!} R^{-2i-j}. \]

This paper has been typeset from a TeX/LaTeX file prepared by the author.

\[ R_s = \left( \frac{6}{n + 5} \right)^{1/2} R, \]
\[ \gamma^2 = \left( \frac{\sigma_1}{\sigma_0 \sigma_2} \right)^2 = \frac{n + 3}{n + 5}, \]
\[ \kappa^2 = \left( \frac{\sigma_2}{\sigma_1 \sigma_3} \right)^2 = \frac{n + 5}{n + 7}, \]
\[ \epsilon = \left( \frac{R_a R_b}{R_h^2} \right)^{a+3/2}, \]
\[ r_1 = \left( \frac{R_1}{R_h} \right)^2, \text{ and } \]
\[ r_2 = \left( \frac{R_1}{R_h} \right)^3. \]