A GENERAL FORMULATION OF THE SOURCE CONFUSION STATISTICS AND APPLICATION TO INFRARED GALAXY SURVEYS

TSUTOMU T. TAKEUCHI$^{1,3}$

AND

TAKAKO T. ISHI$^{2,3}$

Draft version January 2, 2005

ABSTRACT

Source confusion has been a long-standing problem in the astronomical history. In the previous formulation of the confusion problem, sources are assumed to be distributed homogeneously on the sky. This fundamental assumption is, however, not realistic in many applications. In this work, by making use of the point field theory, we derive general analytic formulae for the confusion problems with arbitrary distribution and correlation functions. As a typical example, we apply these new formulae to the source confusion of infrared galaxies. We first calculate the confusion statistics for power-law number counts as a test case. When the slope of differential number counts, $\gamma$, is steep, the confusion limits becomes much brighter and the probability distribution function (PDF) of the fluctuation field is strongly distorted. Then we estimate the PDF and confusion limits based on the realistic number count model for infrared galaxies. The gradual flattening of the slope of the source counts makes the clustering effect rather mild. Clustering effects result in an increase of the limiting flux density with $\sim 10\%$. In this case, the peak probability of the PDF decreases up to $\sim 15\%$ and its tail becomes heavier. Though the effects are relatively small, they will be strong enough to affect the estimation of galaxy evolution from number count or fluctuation statistics. We also comment on future submillimeter observations.

Subject headings: methods: statistical — galaxies: statistics — infrared: galaxies — large-scale structure of universe — submillimeter — surveys

1. INTRODUCTION

Astronomical source counts often suffer from a problem that multiple sources are located in a single beam of the observational instrument. The obtained number counts will then be different from true ones, because the apparent position and flux of the sources are changed by blending of other, usually fainter sources. This is called the source confusion, and the resulting measurement error is referred to as the confusion noise.

In its early stage, confusion problem has been investigated and formulated mainly by radio astronomers in relation to the so-called fluctuation analysis or $P(D)$ analysis (???????). This analysis has been immediately applied and developed further in X-ray (??, e.g.,) scheuer74,franceschini82,barcons90,barcons92,barcons94 and infrared astronomy (??, e.g.,) hacking87,franceschini89,oliver97,matsuura00,lagache00, miville02,friedmann03. Now confusion problem also becomes important in submillimeter cosmology and high-precision astrometric missions (??, cf.) hogg01.

Cosmologists compare model predictions with observed source counts to extract information of the evolution of galaxies and other objects (??, e.g.,) franceschini91,guiderdoni98, takeuchi99,hirashita99,franceschini01,takeuchi01a,takeuchi01b. The confusion noise often prevents us from direct comparison between measured counts and predictions. Especially, recent advent of infrared (IR) and submillimeter facilities may have stimulated the discussion on the confusion problem. Compared with other wavelengths, we have relatively small aperture telescopes in the IR, mainly because ground-based observations are almost impossible, and we inevitably need cooled space telescopes. In the submillimeter the antenna is relatively large up to $\sim 10$–$30$ m, but the long wavelength also makes the single-dish surveys confusion-limited by a large diffraction.

Generally, the measurement error of the source flux blurs the true source counts. This problem was originally pointed out by ??) and deeply considered in subsequent studies (??, e.g.,) eddington40,bennett62,refsdal69,murdoch73,hogg98, but there still remains some unsolved issues, especially when the error is dominated by the confusion noise. Recently Monte Carlo simulation has become popular to evaluate confusion effect (??, e.g.,) bertin97,eales00,scott02, but such an approach is sometimes not easy to interpret, and often time consuming. Today, many large survey projects will be performed or completed soon, and a general analytical prescription for the confusion is desirable.

Fundamental assumption for the formulation of the confusion problem is that the sources are distributed homogeneously on the sky (??, e.g.,) scheuer57,condon74,franceschini89. However, this assumption is obviously not realistic in many applications, e.g., stars in the Galaxy, or galaxies in the Universe. A straightforward attempt to take into account the source clustering is to integrate over the spatial correlation functions along the line of sight. This approach has been adopted by several works (??, e.g.,) barcons88,barcons98,bugian98. Among them, ??) presented comprehensive results closely related to the present issue. Since it requires the knowledge of three-dimensional correlation functions, while it is

$^1$National Astronomical Observatory of Japan, Mitaka, Tokyo 181–8588, JAPAN; takeuchi@optik.mtk.nao.ac.jp.

$^2$Kwasan and Hida Observatories, Kyoto University, Yamashina-ku, Kyoto 607–8471, JAPAN; ishii@kwasan.kyoto-u.ac.jp.

$^3$Research Fellows of the Japan Society for the Promotion of Science.
a convenient method for theoretical predictions, observationally it is quite rare that we can obtain spatial information at the early phase of a survey. Hence, we need a method to evaluate the confusion only from the projected information of source clustering. In this line, \( ? \) has pioneered the methodology to tackle the riddle, but his results were restricted to a few simple cases. Since then, little analytical progress on this topic has been made up to now.

In this work, by making extensive use of the methods of the theory of point process, we show the general analytic formula for the confusion problems with arbitrary distribution and correlation functions. In Section 2, we first consider the fluctuation of unclustered sources, and then present the general formulae for inhomogeneous and clustered source distributions. Based on the new formulae, we reconsider how to treat the confusion noise in Section 3. In Section 4 we formulate higher-order clustering of galaxies, which will be included in the general confusion problems. We focus on far-infrared and submillimeter galaxies. Section ?? is devoted to our conclusions. Factorial moments and cumulants are introduced in Appendix ??, and then we translate it to the language of the theory of point fields (\( ? \)). By this method, we can straightforwardly extend our formulation to the confusion problem of clustered sources (\( ?, \text{cf.} \))\[\text{barcons92}].

2. STATISTICS OF THE FLUCTUATION FIELD

Statistics of fluctuation field is a fundamental tool for the confusion problem. We first formulate the fluctuation of confusion noise in line with classical works (\( ? \))\[\text{scheuer57},\text{scheuer74} \], and then we translate it to the language of the theory of point fields (\( ? \)). By this method, we can straightforwardly extend our formulation to the confusion problem of clustered sources (\( ?, \text{cf.} \))\[\text{barcons92}].

2.1. Fluctuation Generated from Sources without Clustering

2.1.1. Classical derivation

We define a probability density function (hereafter PDF) of a source having a flux \( S' \in [S, S + dS], p(S) dS \), i.e.,

\[
\int_{0}^{\infty} p(S) dS = 1 .
\]  

We set \( h(x) = h(\theta, \phi) \), the beam pattern (normalized to unity at the beam center), and \( s(x) = Sh(x) \), the system response from the intrinsic flux \( S \).

The total signal response we observe, \( I(x) \), is described as

\[
I(x) = \sum_{n=0}^{\infty} S_n h(x - x_n) .
\]

where \( S_n \) is the flux of the \( n \)-th source, and \( x_n \) is the angular position of the source. For the following discussion we need to define a flux signal that consists of exactly \( N \) sources, \( I_N(x) \),

\[
I_N(x) = \sum_{n=0}^{N} S_n h(x - x_n) .
\]

Consider an ensemble of the number of sources in a unit solid angle, \( n_\ell, \ell = 1, 2, \ldots \), and let the mean of \( n_\ell \) is \( \bar{n} \). Here, \( \mathcal{N}(S) \equiv \bar{n} p(S) \) gives the familiar differential number count of the considered sources. Then, from the no clustering assumption, the probability of observing exactly \( N \) sources in a beam is given by the Poisson process

\[
p_N = \frac{(\bar{n} \Omega_b)^N}{N!} e^{-\bar{n} \Omega_b} .
\]

where \( \Omega_b \) is the solid angle of the beam. Then let us consider the PDF of the value of \( I(x) \), \( f(I) \),

\[
f(I) = \text{Prob} \{ I' \in [I, I + dI] \}
\]

\[
= \sum_{N=0}^{\infty} \text{Prob} \{ \text{exactly } N \text{ sources lie in a beam} \} \text{Prob} \{ I_N(x) \in [I, I + dI] \}
\]

\[
\equiv \sum_{N=0}^{\infty} p_N g_N(I) ,
\]

where \( g_N(I) \) denotes the PDF of \( I_N(t) \).

Under a certain regularity condition, a PDF is uniquely characterized via its Laplace transform (LT)\[\text{\cite{LT}}\]. We define the LTs of \( f(I) \) and \( g_N(I) \), \( \mathcal{L}_f(t) \) and \( \mathcal{L}_g_N(t) \), respectively,

\[
\mathcal{L}_f(t) \equiv \mathbb{E} \left[ e^{-tI} \right] = \int_{-\infty}^{\infty} e^{-tI} f(I) dI ,
\]

\[
\mathcal{L}_g_N(t) = \int_{-\infty}^{\infty} e^{-tI} g_N(I) dI ,
\]

\[\text{4This function corresponds to } P(D) \text{ in radioastronomical terminology.}\]

\[\text{5If we use } it \text{ instead of } t, \text{ we obtain a characteristic function (CF). While the CF is also quite common in physical studies, here we use LT to refer literatures in mathematics and statistics.}\]
where \( E[\cdot] \) represents the expectation value of a random variable. Then we obtain

\[
\mathcal{L}_f(t) = \sum_{N=0}^{\infty} p_N \mathcal{L}_{g_N}(t).
\]

(8)

Hence, concrete expression of \( \mathcal{L}_{g_N}(t) \) is our next step to have the functional form of \( f(I) \).

We consider random variables \( s_n = S_n h(x - x') \) and their summation over \( n \), \( I_N(x) = \sum_{n=0}^{\infty} s_n \). Then the PDF of a signal \( I \) with \( N \) summands is represented by the following recursive convolution:

\[
g_{N+1}(I) = \int_0^I g_N(I - s') g(s') ds',
\]

\[
g_1(I) = g(I).
\]

(9)

(10)

Since the LT of a convolution of two functions is a normal product of the LTs of them, the PDF of \( s \),

\[
g(s) \equiv \text{Prob} \{ s' \in [s, s + ds] \}
\]

(11)

gives

\[
\mathcal{L}_{g_N}(t) = \mathcal{L}_g(t)^N,
\]

(12)

where

\[
\mathcal{L}_g(t) = \int_{-\infty}^{\infty} e^{-ts} g(s) ds.
\]

(13)

Then we have

\[
\mathcal{L}_f(t) = \sum_{N=0}^{\infty} \frac{\langle \bar{n}\Omega \rangle^N e^{-\bar{n}\Omega}}{N!} \mathcal{L}_g(t)^N
\]

\[
= e^{-\bar{n}\Omega} \sum_{N=0}^{\infty} \frac{\langle \bar{n}\Omega \rangle^N \mathcal{L}_g(t)^N}{N!}
\]

\[
= e^{-\bar{n}\Omega} e^{\bar{n}\Omega} \mathcal{L}_g(t) = e^{\bar{n}\Omega} [\mathcal{L}_g(t) - 1].
\]

(14)

The rest we should consider is the exact form of \( \mathcal{L}_g(t) \). The profile of a single source is expressed as \( s(x) = Sh(x) \). Since the LT of a random variable \( s \) is a statistical average of \( e^{-ts} \), we obtain

\[
\mathcal{L}_g(t) = \frac{1}{\Omega_b} \int_{\Omega_b} \int_{S} e^{-tSh(x)} p(S)dSd\mathbf{x}.
\]

(15)

Here we used a simplified symbol \( \int_{S} \equiv \int_{S=0}^{\infty} \); i.e., integration with respect to \( S \) over the range of \([0, \infty] \). We use this notation for some other variables in the following.

Substituting Equation (15) into Equation (14), we finally obtain

\[
f(I) = \mathcal{L}^{-1} \left[ e^{\bar{n}\Omega} [\mathcal{L}_g(t) - 1] \right]
\]

\[
= \mathcal{L}^{-1} \left[ \exp \bar{n} \int_{\Omega_b} \left[ \int_{S} e^{-tSh(x)} p(S)dS - 1 \right] d\mathbf{x} \right],
\]

(16)

where \( \mathcal{L}^{-1} [\cdot] \) represents the inverse Laplace transform. By the formula for moments

\[
\mu_k \equiv E[I^k] = (-1)^k \left. \frac{d^k \mathcal{L}_f(t)}{dt^k} \right|_{t=0},
\]

(17)

and cumulants (reduced moments)

\[
\kappa_k = (-1)^k \left. \frac{d^k \ln \mathcal{L}_f(t)}{dt^k} \right|_{t=0},
\]

(18)

(?, e.g.,) we obtain the following beautiful expressions for the cumulants of \( I \) (?, the simplest result of Campbell’s theorem:)|campbell09,rice44.

\[
\kappa_1 = E[I] = \mu_1 = \bar{n} \int_{\Omega_b} \int_{S} Sh(x)p(S) dSd\mathbf{x} = \int_{\Omega_b} \int_{S} Sh(x)\mathcal{M}(S) dSd\mathbf{x},
\]

(19)

\[
\kappa_2 = E[(I - \mu_1)^2] = \int_{\Omega_b} \int_{S} S^2 h(x)^2 \mathcal{M}(S) dSd\mathbf{x}, \ldots,
\]

(20)

and generally,

\[
\kappa_k = \int_{\Omega_b} \int_{S} S^k h(x)^k \mathcal{M}(S) dSd\mathbf{x}.
\]

(21)
2.1.2. Alternative derivation via point field theory

The position of galaxies, stars, or other point sources is expressed as a point in the field under consideration. A mathematical technique to treat such a field of points is called the theory of point process, and has a long history (?, e.g.,)stoyan94,stoyan95,daley03. The theory provides us very powerful tools for the problems that we consider here. Thus, we translate the above heuristic derivation of \( f(I) \) with the language of point process theory. Considering the problem along this line enables us a straightforward extension of our discussion to the case of clustered sources.

The fluctuation field \( I(x) \) is expressed as

\[
I(x) = \sum_{n=0}^{\infty} S_n h(x - x_n)
\]

\[
= \int_{\mathbb{R}^2} h(x - x') S(x') \mathcal{N}(dx') ,
\]

where \( \mathcal{N}(A) \) of Borel sets \( A \in \mathbb{R}^2 \) is the so-called ‘counting measure’, which represents the number of points in the set \( A \), and \( S(x') \) is a fictitious stochastic process that takes a value \( S(x'_n) = S_n \) at each point \( x'_n \) (?). Here we identify the celestial sphere with a real plane \( \mathbb{R}^2 \). We define a random measure \( \mathcal{M}(A) \) as

\[
\mathcal{M}(A) \equiv \int_A I(x) dx
\]

\[
= \int_A \left[ \int_{\mathbb{R}^2} h(x - x') S(x') \mathcal{N}(dx') \right] dx
\]

\[
= \sum_{x_n \in A} S_n \int_{\mathbb{R}^2} h(x - x_n) dx .
\]

Here we introduce a Laplace functional \( \mathcal{L}_M[\mathcal{X}] \) with respect to the random measure \( \mathcal{M} \),

\[
\mathcal{L}_M[\mathcal{X}] \equiv \mathbb{E} \left[ \exp \left[ - \int_{\mathbb{R}^2} \mathcal{X}(x) \mathcal{M}(dx) \right] \right]
\]

\[
= \mathbb{E} \left[ \exp \left[ - \int_{\mathbb{R}^2} \mathcal{X}(x) I(x) dx \right] \right] .
\]

It is a Laplace counterpart of the characteristic functional of a random field, both of which are often used in particle physics, turbulence theory, and structure formation theories in the Universe, etc. (?), e.g.,)vlad94,frisch95,szapudi93,matsubara95. We observe

\[
\int_{\mathbb{R}^2} \mathcal{X}(x) \mathcal{M}(dx) = \int_{\mathbb{R}^2} \mathcal{X}(x) \left[ \sum_{n} S_n h(x - x_n) \right] dx
\]

\[
= \sum_{n} S_n \int_{\mathbb{R}^2} \mathcal{X}(x) h(x - x_n) dx .
\]

Here,

\[
\mathbb{E} \left[ \exp \left[ - S_n \int_{\mathbb{R}^2} \mathcal{X}(x) h(x - x_n) dx \right] \right] = \mathcal{L}_S \left[ \int_{\mathbb{R}^2} \mathcal{X}(x) h(x - x_n) dx \right] \equiv \mathcal{Z}(x_n) .
\]

In Equation (26), \( \mathcal{L}_S[\cdot] \) represents the Laplace-Stieltjes functional with respect to the PDF of \( S_n \), \( p(S_n) \), i.e.,

\[
\mathcal{L}_S \left[ \int_{\mathbb{R}^2} \mathcal{X}(x) h(x - x_n) dx \right] \equiv \int_S \exp \left[ - S_n \int_{\mathbb{R}^2} \mathcal{X}(x) h(x - x_n) dx \right] p(S_n) dS_n .
\]

Combining Equations (24), (25), and (26), we have an important relation

\[
\mathcal{L}_M[\mathcal{X}] = \prod_n \mathbb{E} \left[ \exp \left[ - S_n \int_{\mathbb{R}^2} \mathcal{X}(x) h(x - x_n) dx \right] \right]
\]

\[
= \prod_n \mathcal{Z}(x_n)
\]

\[
= G[\mathcal{Z}] .
\]

The last step is the definition of the probability generating functional (PGFL) of a point field, \( G[\mathcal{Z}] \) (?), cf.)balian89,daley03. This shows an important fact: the fluctuation field blurred by a beam profile is expressed in terms of the PGFL of the

\[
\text{In statistical and particle physics, } \mathcal{L}_M[\mathcal{X}] \text{ is often expressed in the form}
\]

\[
\mathcal{L}_M[\mathcal{X}] = \int \mathcal{D}[I] \mathbb{E} \left[ \mathcal{X}(x) I(x) dx \right]
\]

in relation to the path integral formulation (?), e.g.,)fry84a,szapudi93,matsubara95. Here \( \mathcal{D}[I] \) is the assigned probability functional for an overall configuration of the field \( I(x) \).
original point field. To be exact, we can describe the observed fluctuation field as an explicit functional of the true point field.

In general, \( G[Z] \) has some useful expansions with respect to the factorial moments, factorial cumulants, and other related summary statistics of the point field \( (\frac{1}{n!}) \) \cite{maa85, vlad94, kerscher01, daley03}. The most familiar statistic for astrophysical studies may be the correlation function (or equivalently, normalized factorial cumulants). Then, \( G[Z] \) can be expressed in the following form \( (\frac{1}{n!}) \) \cite{maa85}:

\[
\ln G[Y + 1] = \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{R}^2 \times \cdots \times \mathbb{R}^2} Y(x_1) \cdots Y(x_k) c_{[k]}(x_1 \cdots x_k) dx_1 \cdots dx_k
\]

\[
= \sum_{k=1}^{\infty} \frac{\bar{n}^k}{k!} \int_{\mathbb{R}^2 \times \cdots \times \mathbb{R}^2} Y(x_1) \cdots Y(x_k) w_k dx_1 \cdots dx_k,
\]

(29)

where \( c_{[k]} \) denotes the factorial cumulant of the point field (see Appendix ??), and \( w_k = w_k(x_1, \cdots, x_k) \) is the angular \( k \)-point correlation function of the point sources. This relation will be used to extend our formulation to the clustered point sources in \( \S 2.3 \). For a Poisson field, by its definition, all the higher order \( (k \geq 2) \) factorial cumulants vanish, and we obtain

\[
\ln G[Y + 1] = \bar{n} \int_{\mathbb{R}^2} Y(x) dx.
\]

(30)

Hence, by substituting \( Z = Y + 1 \) in Equation (30) and doing some algebra, we obtain

\[
G[Z] = e^{\bar{n} \int_{\mathbb{R}^2} [Z(x) - 1] dx}.
\]

(31)

Now returning back to Equation (24),

\[
\mathcal{L}_M[\chi] = G[Z] = e^{\bar{n} \int_{\mathbb{R}^2} [Z(x) - 1] dx}.
\]

(32)

In order to obtain the formula for the local process in a beam area \( \Omega_b \), we set the test function \( Z(x) \) as

\[
Z^*(x) = 1 - [1 - Z(x)] \mathbb{I}_{\Omega_b},
\]

(33)

where \( \mathbb{I}_A \) is an indicator function of a set \( A \),

\[
\mathbb{I}_A = \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases}
\]

(34)

Substituting Equation (33) into Equation (32) yields

\[
\mathcal{L}_M[\chi] = e^{\bar{n} \int_{\mathbb{R}^2} [Z^*(x) - 1] dx} = e^{\bar{n} \int_{\mathbb{R}^2} [Z(x) - 1] \mathbb{I}_{\Omega_b} dx} = e^{\bar{n} \int_{\Omega_b} [Z(x) - 1] dx}.
\]

(35)

Recall that

\[
Z(x_n) = \mathcal{L}_M \left[ \int_{\mathbb{R}^2} \chi(x) h(x - x_n) dx \right],
\]

(36)

we obtain

\[
\mathcal{L}_M[\chi] = e^{\bar{n} \int_{\Omega_b} \left\{ \mathcal{L}_M \left[ \int_{\mathbb{R}^2} \chi(x') h(x' - x) dx' \right] - 1 \right\} dx} = e^{\bar{n} \int_{\Omega_b} \left\{ \int_S e^{-S \int_{R^2} \chi(x') h(x' - x) dx'} p(S) dS - 1 \right\} dx}.
\]

(37)

By setting the test function \( \chi(x') = t \), we have

\[
\mathcal{L}_f(t) = e^{\bar{n} \int_{\Omega_b} \int_S e^{-tS h(x' - x)} p(S) dS - 1} dx.
\]

(38)

Inverting the Laplace transform, we again obtain the desired result

\[
f(I) = \mathcal{L}^{-1} \left[ e^{\bar{n} \int_{\Omega_b} \int_S e^{-tS h(x) p(S) dS - 1} dx} \right],
\]

(39)

which is the same result with Equation (16).
2.2. The Case of Inhomogeneous Poisson Point Field

Sometimes we face a problem that the points distribute locally Poisson but the intensity \( m \) is spatially inhomogeneous, i.e., depends on the position on the sky. The projected stellar density distribution of the Milky Way might be described in this way (c.f., Chandrasekhar50).

If the typical angular scale of the variation of \( \bar{n} \) is larger than the typical beam size, we only have to divide the sky into patches with the variation scale and derive the PDF of the fluctuation \( f(I) \) in each patch by Equation (16).

On the other hand, the variation scale of \( \bar{n} \) is comparable or smaller than the beam size, we should properly treat the variation within a beam. In such case \( \bar{n} \) is expressed as \( \bar{n}(x) \) at the scale of our interest. The intensity in a set \( A \), \( \bar{N}(A) \), is obtained by integrating \( \bar{n}(x) \) over \( A \), as

\[
\bar{N}(A) = \int_A \bar{n}(x) dx,
\]

(7) pp.650–652, cressie93. In this case the joint probability for some positions are still Poisson, and the independence still holds. It leads to the following formula

\[
\mathcal{L}_f(t) = \exp \int_{\Omega_b} \left[ \int_S e^{-t \bar{S}(x)} p(S) dS - 1 \right] \bar{n}(x) dx.
\]

The same as the above, the PDF \( f(I) \) is obtained by inverting the LT.

2.3. Fluctuation Generated from Clustered Sources

We now turn to the case of the point source with significant clustering. Clustering of the sources makes the confusion effect more severe, because the variance of the source number is larger than that of Poisson fluctuation. The importance of clustering in the fluctuation and confusion problems has already pointed out as early as the end of 1960’s by (7), but he did not provided a quantitative answer in that paper. The fluctuation analysis of sky brightness including the effect of clustering was first presented by (7) for the study of the X-ray background. The central tool for his analysis was characteristic functional of the field \( I(x) \). We use its equivalent, Laplace functional here, and derive the formula in mathematically rigorous way, in parallel with the discussion in §2.2.1.

Again we start with the fluctuation field \( I(x) = \sum_{n=0}^{\infty} S_n h(x - x_n) \), but this time \( x_n \) are not distributed at random on the sky, but have a certain correlation with each other. The statistical properties of the field are characterized by \( \mathcal{L}_M[I] \). Since the derivation of Equation (28) does not depend on the clustering property of the source point field, we can also apply Equations (28) and (29). Since the point field has non-vanishing correlation functions for clustered sources, Equation (29) reads

\[
\ln G[Z] = \sum_{k=1}^{\infty} \frac{\bar{n}^k}{k!} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} [Z(x_1) - 1] \cdots [Z(x_k) - 1] w_k(x_1, \ldots, x_k) dx_1 \cdots dx_k,
\]

therefore

\[
G[Z] = \exp \sum_{k=1}^{\infty} \frac{\bar{n}^k}{k!} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} [Z(x_1) - 1] \cdots [Z(x_k) - 1] w_k(x_1, \ldots, x_k) dx_1 \cdots dx_k.
\]

The same as the unclustered field, we set \( Z^* = 1 - (1 - Z) \bar{z}_{\Omega_b} \), and obtain

\[
G[Z] = \exp \sum_{k=1}^{\infty} \frac{\bar{n}^k}{k!} \int_{\Omega_b} \cdots \int_{\Omega_b} [Z(x_1) - 1] \cdots [Z(x_k) - 1] w_k(x_1, \ldots, x_k) dx_1 \cdots dx_k.
\]

Substituting Equation (36) yields

\[
\mathcal{L}_M[I] = \exp \sum_{k=1}^{\infty} \frac{\bar{n}^k}{k!} \int_{\Omega_b} \cdots \int_{\Omega_b} \left[ \int_{\mathbb{R}^2} \mathcal{L}_M[I(x) h(x - x_j) dx] - 1 \right] w_k(x_1, \ldots, x_k) dx_1 \cdots dx_k.
\]

Again by letting \( \mathcal{X}(x) = t \), we have

\[
\mathcal{L}_f(t) = \exp \sum_{k=1}^{\infty} \frac{\bar{n}^k}{k!} \int_{\Omega_b} \cdots \int_{\Omega_b} \left[ \int_{S_j} e^{-t S_h(x - x_j)} p(S_j) dS_j - 1 \right] w_k(x_1, \ldots, x_k) dx_1 \cdots dx_k
\]

\[
= \exp \sum_{k=1}^{\infty} \frac{\bar{n}^k}{k!} \int_{\Omega_b} \cdots \int_{\Omega_b} \left[ \int_{S_j} e^{-t S_h(x)} - 1 \right] p(S_j) dS_j w_k(x_1, \ldots, x_k) dx_1 \cdots dx_k.
\]

Laplace inversion gives the final general formula for the PDF of intensity fluctuation, \( f(I) \).
Fig. 1.—A schematic description of the confusion noise problem. Top panel: The source with flux $S$ is embedded in the fluctuation field. Middle panel: The source is observed with flux $S + I$ if the total noise intensity in the beam is $I$. Bottom panel: the source is observed as an additive noise of the brighter source with flux $S' > S$ if $I$ is dominated by one or more bright sources $S' > S$ (dashed line). On the other hand, if $I$ consists of some sources fainter than $S$, $I$ should be merely added to $S$ as a noise even if it is larger than $S$.

We obtain cumulants of the PDF of $I$ by the same procedure as Equations (19)–(20):

$$\kappa_1 = \int_{\Omega_b} \int_S S h(x) N(S) dS dx,$$

$$\kappa_2 = \int_{\Omega_b} \int_S S^2 h(x)^2 N(S) dS dx + \int_{\Omega_b} \int_{\Omega_b} \left[ \prod_{j=1}^2 \int S_j h(x_j) N(S_j) dS_j \right] w_2 dx_1 dx_2,$$

and so on (see Appendix ??). It is a natural result that $\kappa_1$ is the same as that of unclustered sources. On the other hand, $\kappa_2$ involves an additional term with respect to $w_2 = w_2(x_1, x_2)$, which causes the fluctuation field overdispersed compared to that of unclustered sources.

3. RECONSIDERATION OF THE CONFUSION PROBLEM

The concept of confusion noise and confusion limit is closely related to the fluctuation field $f(I)$. ? formulated the confusion limit for a power-law source number counts. His formulation was then extended for general number counts by ?). ? have made some improvement for precise calculation by the method. We should note that the confusion noise affects source counts up to the flux much larger than that of confusion limit (?). We reconsider the confusion problems in this section.

3.1. Confusion Noise

First we review the problems related to the confusion noise. ? proposed an important concept that existence of any statistical noise makes the observed source counts biased upward. Now this well-known phenomenon is called the ‘Eddington bias’. He considered the case that the noise is Gaussian. After that, a large amount of efforts has been made to treat and correct the bias by subsequent authors (e.g.,) eddington40,bennett62,murdoch73. The outline of these ideas is briefly summarized in ?), in comparison with their own method.
Problems with Gaussian noise is beautifully formulated by?). Their setting assumes that the dispersion of the error does not depend on the source flux. Even in such a simple case, (?) showed that the confusion noise can be very severe for sources with low signal-to-noise ratios. They pointed out that when the confusion noise dominate the error, the total noise cannot be Gaussian, and gave a qualitative discussion on the problem. They claim to use a Monte Carlo approach to evaluate in that case.

However, as we will discuss in the subsequent subsections, we have an expression for the fluctuation field, and we can utilize the information for evaluating the effect of the confusion noise in general cases. (?) properly distinguish between the problems of confusion and blending: the former generally means that the source flux is affected by fainter sources, which cannot be detected by the considered instrument individually, and the latter indicates that the source flux is affected by fainter sources, which is bright enough to be detected as a source if they were located far from the central brightest source. The term ‘confusion limit’ is related to the former phenomenon, and consequently, confusion limit flux is weaker than the noise flux for bright sources.

How to treat these two effects in a unified manner? Of course they occur in a same way, hence there is no distinction in principle. Consider an ideal situation that there are no detector noise, photon noise, read-out noise, nor any other cosmological diffuse background radiation with structure. Then the noise intensity distribution is the fluctuation PDF itself. Here we consider a source with true flux \( S \). The situation is schematically described in Figure 1. The source is inevitably affected by other sources in a same beam. The total noise intensity in the beam is \( I \). If \( I < S \), then the source is observed with flux \( S + I \) (see the middle panel of Figure 1). On the other hand, if \( I \) is dominated by a bright source \( S' > S \), then the source is observed as an additive noise of the brighter source with flux \( S' > S \) (the bottom panel of Figure 1). However, (?) pointed out that there is a possibility that the signal \( I \) consists of two or more sources with fluxes fainter than \( S \). The probability of such multiple source blending is obtained by the same manner as discussed in \$2\). Again we should consider \( s_j = S_j h(x - x_j) \), under the condition

\[
\sum_{j=1}^{k} s_j = I,
\]

and the probability of having a noise intensity \( I \) produced by \( k \) sources is expressed by a convolution [Equation (9)]. For \( I \) to be a mere noise, there must not be sources brighter than \( S \) in a beam. Hence, we should derive the conditional probability of having \( I \) caused by \( k \) sources under the condition \( S_j < S, j = 1, \ldots, k \), and then sum up for all \( k \)s. It may seem to be a complicated problem, but it can be simplified by Laplace transform. By restricting the range of integration from 0 to \( S \) in Equation (15) we have the LT of the conditional PDF as

\[
\mathcal{L}_{\phi_b}(t; S) = \left[ \frac{1}{\Omega_b} \int_{\Omega_b} \int_{0}^{S} e^{-tS'h(x)} p(S') dS' dx \right]^k.
\]  

(49)

Then we have the LT of the desired total conditional PDF, \( \tilde{f}(I; S) \), as

\[
\mathcal{L}_{\tilde{f}}(t; S) = \exp \left[ \frac{\bar{n}}{\Omega_b} \int_{\Omega_b} \left[ \int_{0}^{S} e^{-tS'h(x)} p(S') dS' \right] dx - 1 \right]
\]

(50)

for unclustered point sources, and

\[
\mathcal{L}_{\tilde{f}}(t; S) = \exp \sum_{k=1}^{\infty} \frac{\bar{n}^k}{k!} \int_{\Omega_b} \cdots \int_{\Omega_b} \prod_{j=1}^{k} \left[ \int_{0}^{S} e^{-tS_j'h(x_j)} p(S_j') dS_j' - 1 \right] w_k \, dx_1 \cdots dx_k
\]

(51)

for clustered sources.

Hence, in order to obtain the confused number counts from the true ones, we should convolve the fluctuation distribution \( f(I) \) itself at \( I < S \), and \( \tilde{f}(I; S) \) at \( I > S \). This consideration naturally explains the fact that the confusion noise flux is stronger than the confusion limit flux. Since we have a variety of background in the image data in a real situation, we fit and subtract from the obtained flux, the noise is not necessarily positive. Thus in practice, the dependence of the confusion noise distribution on flux \( S \) causes only a weak variation along \( S \) under the existence of other kind of noise, because of the convolution with other noise.

### 3.2. Confusion limit

We next formulate the relation between the beam size \( \theta_b \) and the source confusion limit\(^7\). The basic procedure is to estimate the limit signal strength as an upper cutoff of the integration in the second-order cumulant formulae in Equation (20) or Equation (48) by iterations. Here we express the second-order cumulant formulae as a function of signal strength \( s \):

\[
\kappa_2 = \int_{\Omega_b} \int_{S} S^2 h(x)^2 \mathcal{M}(S) dS dx
\]

\[
= \int_{\Omega_b} S^2 \int_{h(x)}^{\infty} \mathcal{M} \left[ \frac{s}{h(x)} \right] \frac{dx}{h(x)} ds,
\]

(52)

\(^7\)Hereinafter we assume that \( h(x) = (\theta, \phi) = h(\theta) \) for simplicity. This is not an essential assumption for the subsequent discussions.
In general, we cannot solve Equation (60) analytically, hence we should calculate it numerically.

for clustered sources, and

\[
\kappa_2 = \int_{\Omega_0} \int_{s} S^2 h(x)^2 \mathcal{N}(S) dSdx + \int_{\Omega_0} \int_{s_1} S_1 S_2 h(x_1) h(x_2) \mathcal{N}(S_1) \mathcal{N}(S_2) w_2(x_1, x_2) dS_1 dS_2 dx_1 dx_2
\]

\[
= \int_{s} \int_{\Omega_0} \mathcal{N} \left[ \frac{s}{h(x)} \right] \frac{dx}{h(x)} ds + \int_{s_1} \int_{s_2} s_1 s_2 \int_{\Omega_0} \mathcal{N} \left[ \frac{s_1}{h(x_1)} \right] \mathcal{N} \left[ \frac{s_2}{h(x_2)} \right] w_2 \frac{dx_1}{h(x_1)} \frac{dx_2}{h(x_2)} ds_1 ds_2.
\]

(53)

for clustered sources. We utilize these formulae for deriving confusion limits in the following.

3.2.1. Confusion Limit for power-law number counts

We begin our discussion with the case that the number count is described by a power-law, according to \(7\):

\[
\mathcal{N}(S) = \alpha S^{-\gamma}.
\]

(54)

For unclustered sources, we obtain the confusion limit flux to a cutoff signal \(s_c\) from Equation (52) as

\[
\sigma(s_c)^2 = \int_{0}^{s_c} \int_{\Omega_0} \mathcal{N} \left[ \frac{s}{h(x)} \right] \frac{dx}{h(x)} ds,
\]

\[
= \int_{0}^{s_c} \int_{\Omega_0} \alpha h(\theta)^{-1} s^{-\gamma} \theta d\theta d\phi
\]

\[
= \int_{0}^{s_c} \alpha \Omega_{\text{eff}} s^{2-\gamma} ds = \left( \frac{\alpha \Omega_{\text{eff}}}{3 - \gamma} \right) s_c^{3-\gamma},
\]

(55)

where \(\Omega_{\text{eff}}\) is the so-called effective beam size, defined as

\[
\Omega_{\text{eff}} = \int_{\Omega_0} h(\theta)^{-1} \theta d\theta d\phi.
\]

(56)

Taking the square root of Equation (55) and setting \(s_c = \alpha \sigma\) as often used, we have

\[
\sigma = \left( \frac{\alpha^{3-\gamma}}{3 - \gamma} \right)^{1/(\gamma-1)} \left(\frac{\pi \theta_{\text{b}}^2}{4 \ln 2 (\gamma - 1)}\right)^{1/(\gamma-1)}.
\]

(57)

If the beam pattern is described by a Gaussian

\[
h(\theta) = e^{-(4 \ln 2)^2 (\frac{\theta}{\theta_{\text{b}}})^2},
\]

(58)

where \(\theta_{\text{b}}\) is the FWHM of the beam, we get Condon's analytic formula for the \(a-\sigma\) confusion limit,

\[
\sigma = \left( \frac{\alpha^{3-\gamma}}{3 - \gamma} \right)^{1/(\gamma-1)} \left[\frac{\pi \theta_{\text{b}}^2}{4 \ln 2 (\gamma - 1)}\right]^{1/(\gamma-1)}.
\]

(59)

It is convenient to relate these expressions and the empirical rule of thumb for the confusion limit. We discuss the relation between the above formulae and the empirical rule of thumb in Appendix ??.

Now we turn to the same problem for clustered sources. For clustered sources, the confusion limit can be obtained in similar way by substituting Equation (54) as follows:

\[
\sigma(s_c)^2 = \int_{0}^{s_c} \int_{\Omega_0} \mathcal{N} \left[ \frac{s}{h(x)} \right] \frac{dx}{h(x)} ds + \int_{0}^{s_c} \int_{\Omega_0} \mathcal{N} \left[ \frac{s_1}{h(x_1)} \right] \mathcal{N} \left[ \frac{s_2}{h(x_2)} \right] w_2 \frac{dx_1}{h(x_1)} \frac{dx_2}{h(x_2)} ds_1 ds_2
\]

\[
= \int_{0}^{s_c} \int_{\Omega_0} \alpha h(\theta)^{-1} s^{-\gamma} \theta d\theta d\phi + \int_{0}^{s_c} \int_{\Omega_0} \alpha^2 h(\theta_1) h(\theta_2) \theta_1 \theta_2 \theta_1 d\theta_1 d\phi_1 \theta_2 d\theta_2 d\phi_2 ds_1 ds_2
\]

\[
= \Omega_{\text{eff}} \left( \frac{\alpha}{3 - \gamma} \right) s_c^{3-\gamma} + \langle \Omega_{\text{eff}}^2 \rangle \left( \frac{\alpha}{2 - \gamma} \right)^2 s_c^{2(2-\gamma)},
\]

(60)

where we defined the following quantity

\[
\langle \Omega_{\text{eff}}^2 \rangle = \int_{\Omega_0} \int_{\Omega_0} [h(\theta_1) h(\theta_2)]^{-1} w_2 \theta_1 \theta_2 \phi_1 \phi_2 d\theta_1 d\phi_1 d\theta_2 d\phi_2
\]

(61)

In general, we cannot solve Equation (60) analytically, hence we should calculate it numerically.

\(^{*}\) Here \(\gamma\) is the power-law exponent of differential number counts. This is the same convention with \(7\). Note that \(7\) uses the same character \(\gamma\) as the exponent of cumulative number counts, and consequently the related expressions are apparently different from those in \(7\) in terms of \(\gamma, \gamma\) also use the cumulative count slope for his numerical study of the confusion errors.
If the beam profile is a Gaussian, and some specific functional form is obtained for Equation (53), we have

the random distribution of galaxies on the sky. In order to obtain the confusion limits more precisely, we must take into
topologies of graphs between points (Fig. 2). For the details of the expansion, see ??

relations for angular galaxy correlation on the sky (??). We apply the following hierarchical
functions to some reasonable order are required. Here we focus on the galaxy clustering as a typical example of the
issues. Of course, we can handle any point sources as far as we have some knowledge of their clustering.

The hierarchical Ansatz is often employed for the correlation functions, both from phenomenological basis (??, e.g.,)balian89
and theoretical considerations of BBGKY equations (??, e.g.,)davis77,fryst84b,yano97. The Ansatz claims that the k-point
correlation is described by a product of k−1 two-point correlation functions. These studies consider the three-dimensional
relations, but we can relate them to two-dimensional ones in a simple way (??). We apply the following hierarchical
relations for angular galaxy correlation on the sky (??):

\[ w_k(x_1, \ldots, x_k) = q_k \sum_{r_j: \text{trees}} w_2(x_{r_1}, x_{r_2}) \cdots w_2(x_{r_{k-1}}, x_{r_{k-1}}), \]

where \( q_k \) is a numerical factor to determine the strength of clustering, and the summation \( \sum_{r_j: \text{trees}} \) is taken over all tree topologies of graphs between points (Fig. 2). For the details of the expansion, see ??).

4. APPLICATION: GALAXY CLUSTERING AND CONFUSION

4.1. Hierarchical Ansatz

As we have discussed above, in order to treat the clustering of the sources, a set of prescribed angular correlation
functions to some reasonable order are required. Here we focus on the galaxy clustering as a typical example of the
related issues. Of course, we can handle any point sources as far as we have some knowledge of their clustering.

Property of higher-order galaxy clustering still remains a matter of debate (??, for a comprehensive overview, see)peebles80,
and various models have been advocated in order to describe the correlation function of galaxies. Generally, a hierarchy
of the correlation functions continues to infinite order, hence we should introduce a certain closure relation to cut the
sequence.

\[ \sigma^2(s_c) = \int_0^{s_c} \int_{\Omega_b} \mathcal{N} \left( \frac{s}{h(x)} \right) \frac{dx}{h(x)} ds + \int_0^{s_c} \int_0^{s_c} \mathcal{N}(\eta) ds \int_{\Omega_b} \mathcal{N} \left( \frac{s_1}{h(x_1)} \right) \mathcal{N} \left( \frac{s_2}{h(x_2)} \right) w_2 \frac{dx_1}{h(x_1)} \frac{dx_2}{h(x_2)} ds_1 ds_2. \]

If the beam profile is a Gaussian, and some specific functional form is obtained for \( w_k = w_k(x_1, \ldots, x_k) \), similar numerical
computation can be performed in parallel with Equation (62).

4.2. Correlation of Infrared Galaxies

Now we focus on the distribution of infrared galaxies, to have some insight to the confusion in the forthcoming infrared
and submillimeter galaxy surveys. Today, many large infrared and submillimeter missions are planned or in progress,
and their source confusion limits for galaxies have been calculated for each facility (??, e.g.,)ishii02,dole03. They all assumes
the random distribution of galaxies on the sky. In order to obtain the confusion limits more precisely, we must take into
account the clustering of infrared galaxies properly.

By using the density moment technique, ?? estimated the coefficients \( q_1-q_8 \) in Equation (66) for IRAS 1.2 Jy sample.
We use these values up to \( k = 4 \) to approximate the clustering: \( q_3 = 1.25 \) and \( q_4 = 2.19 \). The two-point angular correlation
function is

\[ w(x_1, x_2) = w(|x_1 - x_2|) = w(\theta_{12}) = \left( \frac{\theta_{12}}{\theta_0} \right)^{-\beta}, \]

3.2.2. Confusion Limit for general number counts

?? presented an iterative formula for the confusion limit of the general number counts. Here we see the result for a
Gaussian beam [Equation (58)]:

\[ \sigma^2(s_c) = \int_0^{s_c} \int_{\Omega_b} \mathcal{N} \left( \frac{s}{h(x)} \right) \frac{dx}{h(x)} ds = \frac{\pi \theta_0^2}{4 \ln 2} \int_0^{s_c} \int_0^{s_c} \mathcal{N}(\eta) ds \mathcal{N}(\eta_b) \]
and $\beta = 0.79$, $\theta_0 = 0.36$ for the IRAS 1.2 Jy sample. We can obtain the clustering of any deeper surveys via scaling relations of the correlation functions with the characteristic depth of the survey, $d_s$ (7):

$$w_2(\theta_{12}) = d_s^{-1} w_2^0(\theta_{12}) ,$$

(68)

$$w_3(\theta_{12}, \theta_{23}, \theta_{31}) = d_s^{-2} w_3^0(\theta_{12}, \theta_{23}, \theta_{31}) ,$$

(69)

$$w_4(\theta_{12}, \theta_{23}, \theta_{34}, \theta_{13}, \theta_{24}) = d_s^{-3} w_4^0(\theta_{12}, \theta_{23}, \theta_{34}, \theta_{13}, \theta_{24}),$$

(70)

where superscript 0 represents that it is evaluated at $S = 1.2$ Jy at $\lambda = 60 \mu m$, i.e., $w_2^0(\theta) = (\theta/0.36)^{-0.79}$, and $d_s$ is the relative characteristic depth of the survey, defined as $d_s \equiv (1.2 \text{[Jy]} / S \text{[Jy]})^{1/2}$. We observe that the contribution of higher-order clustering rapidly decreases with increasing depth of the survey. This dilution of clustering is more effective for higher-order correlations.

Now we obtain the two-point correlation function $w_2(\theta)$ at an arbitrary flux $S$ through Equation (68) as

$$w_2(\theta) = d_s^{-(1+\beta)} w_2^0(\theta) = d_s^{-1.79} w_2^0(\theta) = \left( \frac{1.2 \text{[Jy]}}{S \text{[Jy]}} \right)^{-1.79/2} w_2^0(\theta).$$

(71)

Consequently, from Equations (69) and (70), we have

$$w_3(\theta_{12}, \theta_{23}, \theta_{31}) = d_s^{-2(1+\beta)} w_3^0(\theta_{12}, \theta_{23}, \theta_{31})$$

$$= d_s^{-3.58} w_3^0(\theta_{12}, \theta_{23}, \theta_{31})$$

$$= d_s^{-3.58} q_3 \left[ w_2^0(\theta_{12}) + w_2^0(\theta_{23}) + w_2^0(\theta_{31}) \right]$$

$$= \left( \frac{1.2 \text{[Jy]}}{S \text{[Jy]}} \right)^{-3.58/2} q_3 \left[ w_2^0(\theta_{12}) + w_2^0(\theta_{23}) + w_2^0(\theta_{31}) \right],$$

(72)

and

$$w_4(\theta_{12}, \theta_{23}, \theta_{34}, \theta_{13}, \theta_{24}) = d_s^{-3(1+\beta)} w_4^0(\theta_{12}, \theta_{23}, \theta_{34}, \theta_{13}, \theta_{24})$$

$$= d_s^{-5.37} w_4^0(\theta_{12}, \theta_{23}, \theta_{34}, \theta_{13}, \theta_{24})$$

$$= d_s^{-5.37} q_4 \left[ w_2^0(\theta_{12}) + w_2^0(\theta_{23}) + w_2^0(\theta_{34}) + w_2^0(\theta_{13}) + w_2^0(\theta_{24}) \right]$$

$$= \left( \frac{1.2 \text{[Jy]}}{S \text{[Jy]}} \right)^{-5.37/2} q_4 \left[ w_2^0(\theta_{12}) + w_2^0(\theta_{23}) + w_2^0(\theta_{34}) + w_2^0(\theta_{13}) + w_2^0(\theta_{24}) \right].$$

(73)
Fig. 3.— The power-law number count model as a model of infrared galaxy counts. These model counts are the integrated ones, hence the power-law indices are represented by $\gamma - 1 = 1.5, 2.0, \text{and} 2.5$.

4.3. Results

4.3.1. Power-law number counts

In order to evaluate how strong the correlation structure affects the confusion-related quantities, we first calculate these statistics based on power-law number count models [Equation (54)],

$$N(S) = \alpha S^{-\gamma},$$

with the power-law indices $\gamma = 2.5, 3.0, \text{and} 3.5$, i.e., the indices of the integrated counts are $\gamma - 1 = 1.5, 2.0, \text{and} 2.5$, respectively. Here, consider a survey at $\lambda = 60 \mu\text{m}$. The power-law number counts are normalized so that they have the same counts with the IRAS QMW galaxy survey at a flux $S = 0.9 \text{Jy}$ (?). It is observationally known that the bright end of the counts is well approximated by Euclidean, because cosmological and evolutionary effects are both negligible. Actually, IRAS QMW galaxy counts shows a good fit to the Euclidean slope (slope index $\gamma = 2.5$ in differential counts), and we can safely fix the slope of the model $\gamma$ to be 2.5 above the flux $S > 0.9 \text{Jy}$. We estimated the confusion-related statistics for the cases with and without clustering. We assume a telescope with an aperture of 70 cm with an ideal Airy function as a PSF, and the detection limit of the instrument is 50 mJy.

As discussed above, the clustering of galaxies depends on the detected flux: brighter sources have stronger clustering on the sky, and the clustering gradually becomes weaker toward fainter flux. This makes the exact formulation for the statistical characteristic of the two-dimensional galaxy distribution prohibitively difficult (?), see]bercons92, and unfortunately, an exact mathematical theory to treat this problem has not fully established yet. In order to calculate the confusion limit with clustering, we approximate the clustering of the whole sample galaxies evaluated at a 'fiducial' flux, $S_{\text{fid}}$, instead of the flux-dependent clustering in gradual way. Fluctuation consists not only of the detected sources but also of unresolvable sources fainter than detection limit, in principle, toward infinitesimally faint flux. Therefore, the fiducial flux is fainter than the point source detection limit. We assumed that the fiducial flux is proportional to detection limit. Based on this assumption, we calibrated the fiducial flux empirically so as to reproduce the IRAS confusion limit (? , $\sigma \approx 20 \text{mJy}$)]hacking87,lonsdale89,berlin97. We found that the relation $S_{\text{fid}} = S_{\text{lim}}/40$ can be used to evaluate the average clustering strength of the IRAS galaxy sample. Namely, all the galaxies with $S > S_{\text{fid}}$ are approximated to have the same correlation functions $[w_2(\theta) = d_{\text{fid}}^{-1}p_2(d_{\text{fid}}\theta)$, etc.] and ignore the contribution from the sources with $S < S_{\text{fid}}$. Of course, exact treatment of this part remains to be theoretically improved.

Table 1

<table>
<thead>
<tr>
<th>Index $\gamma - 1$</th>
<th>5-σ Random [mJy]</th>
<th>5-σ Clustering [mJy]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>3.1</td>
<td>3.8</td>
</tr>
<tr>
<td>2.0</td>
<td>22.9</td>
<td>120</td>
</tr>
<tr>
<td>2.5</td>
<td>138</td>
<td>9332</td>
</tr>
</tbody>
</table>
We calculated the confusion limits from the power-law model counts. Without clustering, the confusion limits would be 3.1 mJy, 22.9 mJy, and 138 mJy for the indices $\gamma - 1 = 1.5$, 2.0, and 2.5, respectively. If we take into account the effect of clustering properly, they become 3.8 mJy, 120 mJy, and 9.33 Jy, for $\gamma - 1 = 1.5$, 2.0, and 2.5, respectively (see Table 1). Thus, the confusion-limit flux get larger, and the effect of clustering strongly depends on the count slope. Especially, if the cumulative count slope $\gamma - 1$ exceeds 2.0, its effect will be catastrophic.

We also obtained the PDF of the fluctuation intensity for infrared galaxies under the same assumptions. Figure 4 shows the PDF with the integral number count slope index $\gamma - 1 = 1.5$. The peak probability decreases by clustering. The clustering also result in a broadening of the PDF (\(\delta\)). It also should noted that the probability of finding a very low intensity at a certain position (closely related to 'the void probability') increases when clustering takes place.

The effect of clustering becomes very strong when the slope index approaches $\gamma - 1 = 2.0$, as presented in Figure 5. In this case, we have already found that the clustering makes the confusion limit much shallower than that of unclustered case. It means that the rms of the confusion noise is large. Actually, we found that the broadening of the PDF by clustering is very strong, and result in a large variance of the fluctuation PDF (Fig. 5). In summary, the effect of clustering can be very severe if the number counts are described by a single power-law. The above result shows that the relative contribution of clustering to the total fluctuation compared with Poisson component becomes larger for large $\gamma$.

These are consistent with the results reported by [4] for dusty galaxies at IR wavelengths. As mentioned above, their approach is more model-oriented than the present work, based on the simple model of the spatial two-point correlation function of galaxies. Their central aim was to estimate the power spectrum of the fluctuation in the submillimeter and radio background, and they considered only the two-point correlation, i.e., second-order statistics. Since confusion limit is the second-order quantity, hence both methods can be used, and yield consistent estimates, though our method requires only projected information.

Our formulation is fundamentally based on the PDF, and it is the most suitable method of calculating it from the