A Maximum Likelihood Method for Bidimensional Experimental Distributions and Its Application to the Galaxy Merger Fraction

CARLOS LÓPEZ-SANJUAN, CÉSAR ENRIQUE GARCÍA-DABÓ, AND MARC BALCELLS

Instituto de Astrofísica de Canarias, Calle Vía Láctea s/n, La Laguna, Tenerife, 38200 Spain; clsj@iac.es, enriques.garcia@gtc.iac.es, balcells@iac.es

Received 2008 January 08; accepted 2008 March 09; published 2008 April 30

ABSTRACT. The determination of the galaxy merger fraction of field galaxies using automatic morphological indices and photometric redshifts is affected by several biases if observational errors are not properly treated. Here, we correct these biases using maximum likelihood techniques. The method takes into account the observational errors to statistically recover the real shape of the bidimensional distribution of galaxies in redshift-asymmetry space, needed to infer the redshift evolution of galaxy merger fraction. We test the method with synthetic catalogs and show its applicability limits. The accuracy of the method depends on catalog characteristics such as the number of sources or the experimental error sizes. We show that the maximum likelihood method recovers the real distribution of galaxies in redshift and asymmetry space even when binning is such that bin sizes approach the size of the observational errors. We provide a step-by-step guide to applying maximum likelihood techniques to recover any one- or two-dimensional distribution subject to observational errors.

1. INTRODUCTION

The currently popular hierarchical ΛCDM models are successful at explaining the structure buildup of the cold dark matter component of the Universe (Springel et al. 2005). But such models have difficulties when explaining the evolution of the baryonic component, even with modeling that incorporates star formation, active galactic nuclei, and supernova feedback, and the multiphase nature of the interstellar medium (De Lucia & Blaizot 2007 and references therein). An open question is the role of the galaxy mergers in the formation of today’s galaxies, especially the most massive ellipticals. The observational determination of the merger rate, $f_{\text{gm}}$, and its evolution with redshift provide empirical clues on the amount and the timing of the merger activity. They also constitute key inputs for semi-analytic models of galaxy formation and evolution.

The merger rate, defined as the number density of merger systems at given redshift, depends on the merger time $\tau_m$, which can only be estimated by $N$-body simulations and simplified models (Mihos 1995; Patton et al. 2000; Conselice 2006). On the other hand, the galaxy merger fraction $f_{\text{gm}}$, defined as the number of merger galaxies in a given galaxy sample in a redshift interval, is a direct observational quantity. Many works have determined the galaxy merger fraction, usually parametrized as $f_{\text{gm}} = f_{\text{gm},0}(1 + z)^m$, using different sample selection and methods, like morphological criteria (Conselice 2003; Lavery et al. 2004; Cassata et al. 2005; Lotz et al. 2008; Bridge et al. 2007), kinematic close companions (Patton et al. 2000; Patton et al. 2002; Lin et al. 2004; De Propris et al. 2005), spatial close pairs (Le Fèvre et al. 2000; Bundy et al. 2004; Bridge et al. 2007; Kartaltepe et al. 2007), or correlation function (Bell et al. 2006; Masjedi et al. 2006). In these works the value of the merger index varies in the range $m = 0 \sim 4$. ΛCDM models predict $m \sim 3$ (Kolatt et al. 1999; Governato et al. 1999; Gottlöber et al. 2001).

The morphological criterion for determining the galaxy merger fraction (see Conselice 2003, hereafter C03) is based on the fact that, just after a merger is complete, the galaxy image shows strong geometrical distortions, in particular, asymmetric distortions. Hence, high values in the automatic asymmetry index $A$ are assumed to identify merger systems. Other automatic morphological indices, such as $M_{20}$ and $G$, have also been used to determine the evolution of galaxy merger fraction with redshift (Lotz et al. 2008). The determination of morphological indices, which must be done on images, is affected by surface brightness dimming and K-corrections, so the errors of the indices grow with redshift and are more important for faint galaxies.

In this paper, we present a method based on the maximum likelihood (ML) technique, to handle the effects of the large errors on the determination of the galaxy merger fraction. Galaxy merger fraction determinations using morphological criteria are generally done on large photometric surveys such as AEGIS (Davis 2007), COMBO-17 (Wolf et al. 2003), COSMOS (Scoville et al. 2007), GOODS (Giavalisco et al. 2007), or SWIRE (Lonsdale et al. 2003). We therefore address the effects of errors in the galaxy asymmetry indices as well as errors on the photometric redshifts.

In § 2 we review the maximum likelihood method for determining bidimensional distributions. Its application to the galaxy merger fraction determination is given in § 2.2. These sections have a high mathematic content, and a statistics background is recommended. Then, in § 3 we summarize the simulations made...
to test the general method and how it improves the galaxy merger fraction determination, § 3.7. In § 4 we provide an outline for the application of the ML method to any one- or two-dimensional experimental distribution subject to observational errors. Our conclusions are presented in § 5.

2. METHODOLOGY

Following Conselice (2006), we define the galaxy merger fraction by morphological criteria as

$$f_{gm} = \frac{\kappa N_m}{N_{tot} + (\kappa - 1)N_m},$$

(1)

where $N_m$ is the number of the distorted sources in the sample, classified as the systems with a value in the asymmetry index $A$ higher than a limiting value $A_m$ (see C03 for details), $N_{tot}$ is the total number of sources in the sample, and $\kappa$ is the average number of galaxies that merged to produce the $N_m$ merger systems. We use $\kappa = 2$ throughout this paper.

In order to compute the galaxy merger fraction and its redshift evolution we must know the underlying distribution of the $z$ and $A$ values, that we assume is represented by a bidimensional histogram in redshift and asymmetry space. This bidimensional histogram is defined by the number of sources in each redshift-asymmetry bin. Normalizing to unity, the histogram yields a bidimensional probability distribution defined now by $p_{kl}$, the probability that a source has redshift in bin $k$ and asymmetry in bin $l$. Index $k$ scans the redshift bins of size $\Delta z$ and index $l$ scans the asymmetry bins of size $\Delta A$. In our case we just need two asymmetry bins separated by $A_m$: the $l = 0$ bin represents normal sources and the $l = 1$ bin represents merger systems. Now, the galaxy merger fraction in redshift bin $[z_k, z_{k+1})$ is

$$f_{gm,k} = \frac{2p_{kl}}{p_{k0} + 2p_{kl}}.$$  

(2)

The accuracy with which the $p_{kl}$ can be obtained degrades significantly when photometric redshifts, $z_{\text{phot}}$, are used, and for typical errors of $A$ in deep surveys. This introduces strong biases in the determination of the galaxy merger fraction.

2.1. The Maximum Likelihood Method

The maximum likelihood method (ML method) developed here is based on García-Dabó (2002), who used this technique to determine unbiased luminosity functions. ML methods have been used in a wide range of topics in astrophysics. Arzner et al. (2007) use it to improve the determination of faint X-ray spectra, Sheth (2007) to obtain redshift and luminosity distributions in photometric surveys, Naylor & Jeffries (2006) to fit color-magnitude diagrams, Makarov et al. (2006) to improve distance estimates using Red Giant Branch stars, and Efstathiou (2004) to analyze low cosmic microwave background multipoles from the Wilkinson Microwave Anisotropy Probe. ML methods are based on the estimation of the most probable values of a set of parameters which define the probability distribution that describes an observational sample (Davidson & Mackinnon 1993; Peña 2001).

The general ML method operates as follows. Throughout the paper we denote as $P(a | b)$ the probability to obtain the values $a$, given parameters $b$. With $z$, being the measured values of source $i$ in the data set and $\theta$ the parameters of the underlying multidimensional distribution that we want to estimate, we may express the joined likelihood function as

$$L(x|\theta) \equiv -\ln \left[\prod_i P(x_i|\theta)\right] = -\sum_i \ln[P(x_i|\theta)],$$

(3)

where $P(x_i|\theta)$ is the probability to obtain $x$, for a given $\theta$. If we are able to express $P(x_i|\theta)$ analytically, we can minimize equation (3) to obtain the best estimation of parameters $\theta$, denoted as $\theta_{\text{ML}}$. In our case, $x_i$ are the observed values of $z$ and $A$ for source $i$, $x_i \equiv (z_{\text{obs},i}, A_{\text{obs},i})$, while $\theta \equiv (p_{kl}, \alpha)$, where $p_{kl}$ are the probabilities which we defined in the paragraph previous to equation (2), and $\alpha$ denotes any other fixed parameters.

Sources are assumed to have real redshift and asymmetry values $z_{\text{real},i}$ and $A_{\text{real},i}$ (not affected by observational errors) that define a bidimensional distribution $p_{kl}$ such that

$$P_{\text{2D}}(z_{\text{real},i}, A_{\text{real},i}|p_{kl}) \equiv \left\{p_{kl}, \forall z_k \leq z_{\text{real},i} < z_{k+1}, A_i \leq A_{\text{real},i} < A_{i+1}\right\}.$$  

(4)

Observational errors cause the observed $z_{\text{obs},i}$ and $A_{\text{obs},i}$ to differ from their respective real values $z_{\text{real},i}$ and $A_{\text{real},i}$. The observed $z_{\text{obs},i}$ and $A_{\text{obs},i}$ are assumed to be extracted for a Gaussian distribution with mean $z_{\text{real},i}$ and standard deviation $\sigma_{z_{\text{obs},i}}$,

$$P_{G}(z_{\text{obs},i}|z_{\text{real},i}, \sigma_{z_{\text{obs},i}}) = \frac{1}{\sqrt{2\pi}\sigma_{z_{\text{obs},i}}} \exp \left(-\frac{(z_{\text{obs},i} - z_{\text{real},i})^2}{2\sigma_{z_{\text{obs},i}}^2}\right).$$  

(5)

Similarly, the observed asymmetry values $A_{\text{obs},i}$ are assumed to be extracted from a Gaussian distribution with mean $A_{\text{real},i}$ and standard deviation $\sigma_{A_{\text{obs},i}}$,

$$P_{G}(A_{\text{obs},i}|A_{\text{real},i}, \sigma_{A_{\text{obs},i}}) = \frac{1}{\sqrt{2\pi}\sigma_{A_{\text{obs},i}}} \exp \left(-\frac{(A_{\text{obs},i} - A_{\text{real},i})^2}{2\sigma_{A_{\text{obs},i}}^2}\right).$$  

(6)

While the $z_{\text{phot}}$ errors may not be strictly Gaussian, this is the best analytical approximation of the errors that we can make. We obtain the probability $P(x_i|\theta)$ of each source by the total probability theorem:
distribution must be normalized to unity, so the only unknowns are the probabilities \( p_{kl} \) which we want to estimate. Note also that we integrate over the variables \( z_{\text{real},i} \) and \( A_{\text{real},i} \), so we are not able to estimate them individually, but only the underlying bidimensional distribution \( p_{kl} \) that describes the sample.

In order to ensure that the probabilities \( p_{kl} \) are non-negative, we change variables, \( p_{kl} = e^{\lambda} \); this change keeps our problem analytic. With these new variables and after integrating equation (7), our likelihood function, defined in equation (3), becomes

\[
L(z_{\text{obs},i}, A_{\text{obs},i} | p_{kl}, \sigma_{z_{\text{obs},i}}, \sigma_{A_{\text{obs},i}}) = \sum_i \left[ \ln \left( \sum_k \sum_l e^{\lambda_k} e^{\lambda_l} \text{ERF}(z_{i,k,l}) \text{ERF}(A_{i,k,l}) \right) \right],
\]

(8)

where

\[
\text{ERF}(\eta, i, k) = \text{erf} \left( \frac{\eta_{\text{obs},i} - \eta_{k}}{\sqrt{2}\sigma_{\text{obs},i}} \right) - \text{erf} \left( \frac{\eta_{\text{obs},i} - \eta_{k}}{\sqrt{2}\sigma_{\text{obs},i}} \right),
\]

and \( \text{erf}(x) \) is the error function. In the minimization of equation (8) we have to take into account that the bidimensional distribution must be normalized to unity,

\[
g(p_{kl}) = \sum_k \sum_l e^{\lambda_k} (z_{k+1,l} - z_k)(A_{l+1,i} - A_l) = 0.
\]

We impose this condition to the ML function by method of the Lagrange multipliers. This is a method for finding the extremum of a function of several variables subject to one or more constraints (see, e.g., Marsden & Tromba 1996 for details). In our case we eventually minimize

\[
G(p_{kl}, \lambda) = L(z_{\text{obs},i}, A_{\text{obs},i} | p_{kl}, \sigma_{z_{\text{obs},i}}, \sigma_{A_{\text{obs},i}}) + \lambda g(p_{kl}),
\]

(11)

where \( \lambda \) is an auxiliary variable. Minimizing equation (11) we obtain the best \( p_{kl} \) values, denoted as \( p_{kl}^{\text{ML}} \). The minimization of equation (11) can be performed with any numerical minimization code. We used \textsc{amoeba}, which is based on the commonly used algorithm of Nelder-Mead (Nelder & Mead 1965) and coded in C (Press 1995 pp. 408–412).

At this point we have the probabilities \( p_{kl}^{\text{ML}} \), and now we want to obtain an estimate of the \( p_{kl}^{\text{ML}} \) covariances. We can obtain those via an expansion of the function \( G(p_{kl}, \lambda) \) in Taylor’s series of its variables \( \theta = (p_{kl}^{\text{ML}}, \lambda) \) around the minimization point \( \theta_{\text{ML}} = (p_{kl}^{\text{ML}, \lambda}) \) if the probability distributions of \( p_{kl}^{\text{ML}} \) are Gaussian, which we assume. The previous minimization process made the first \( G \) derivate null and we obtain

\[
G = G(\theta_{\text{ML}}) + \frac{1}{2}(\theta - \theta_{\text{ML}})^T H(\theta - \theta_{\text{ML}}),
\]

(12)

where \( H = h_{xy} \) is the Hessian matrix and \( T \) denotes the transpose matrix. The inverse of the Hessian matrix gives us an estimate of the 68\% confidence intervals of \( p_{kl}^{\text{ML}} \), denoted as \( [p_{kl}^{\text{ML}, -\sigma_{p_{kl}^{\text{ML}}}}, p_{kl}^{\text{ML}, +\sigma_{p_{kl}^{\text{ML}}}}] \), and the covariances between each \( p_{kl}^{\text{ML}} \), denoted as \( \text{cov}(p_{kl}^{\text{ML}, i}, p_{kl}^{\text{ML}, j}) \), because maximum likelihood theory states that \( \text{cov}(\theta_1, \theta_2) \geq h_{xy}^{-11} \) and \( \sigma_{\theta_1} \geq h_{xx}^{-1/2} \).

In our case, the Hessian matrix is

\[
H = \left( \begin{array}{c} \frac{\partial^2 G}{\partial p_{m,n} \partial p_{s,t}} \\ \frac{\partial^2 G}{\partial \lambda \partial p_{m,n}} \end{array} \right)
\]

(13)

where

\[
\frac{\partial^2 G}{\partial p_{m,n} \partial p_{s,t}} = -\sum_i \frac{\text{ERF}(z_{i,m}, \theta) \text{ERF}(A_{i,n}, \theta)}{16} \times \frac{\text{ERF}(z_{i,s}, \theta) \text{ERF}(A_{i,t}, \theta)}{\sum_k \sum_l e^{\lambda_k} \text{ERF}(z_{i,k,l}) \text{ERF}(A_{i,k,l})},
\]

\[
\frac{\partial^2 G}{\partial \lambda \partial p_{m,n}} = -\frac{1}{\sigma_{\theta}^2} (z_{m+1,n} - z_{m})(A_{n+1,m} - A_{n}) e^{\theta}. \]

Finally, the \( p_{kl}^{\text{ML}} \) probabilities simply are:

\[
p_{kl}^{\text{ML}} = e^{\theta_{kl}^{\text{ML}}}. \]

(16)

Assuming that the \( p_{kl}^{\text{ML}} \) follow a Gaussian distribution, which is assured by the ML theory for a large number of sources, the \( p_{kl}^{\text{ML}} \) follow a log-normal distribution:

\[
P_{\text{LN}} \left( p_{kl} | p_{kl}^{\text{ML}}, \sigma_{p_{kl}^{\text{ML}}} \right) = \frac{e^{-\left( \ln p_{kl} - \ln p_{kl}^{\text{ML}} \right)^2/2\sigma_{p_{kl}^{\text{ML}}}}}{\sqrt{2\pi\sigma_{p_{kl}^{\text{ML}}}}},
\]

(17)

which is highly asymmetric and whose 68\% confidence interval is \([\sigma_{p_{kl}^{\text{ML}, -}}, \sigma_{p_{kl}^{\text{ML}, +}}]\), where

\[
\sigma_{p_{kl}^{\text{ML}, -}} = e^{-\sigma_{p_{kl}^{\text{ML}}}} p_{kl}^{\text{ML}},
\]

(18)

\[
\sigma_{p_{kl}^{\text{ML}, +}} = e^{\sigma_{p_{kl}^{\text{ML}}}} p_{kl}^{\text{ML}}.
\]

(19)

Furthermore, each \( p_{kl}^{00} \) and \( p_{kl}^{01} \) are connected by the covariance...
cov \left( p'_{0i,ML}, p'_{1j,ML} \right)\), so the confidence intervals of \( p_{0i} \) and \( p_{1i} \) are not independent. In the next section we explain how to obtain the confidence interval of the galaxy merger fraction taking this into account.

2.2. The Galaxy Merger Fraction

Expressing the galaxy merger fraction in the range \([z_k, z_{k+1}]\) (eq. [1]) as a function of the output variables of the ML method we obtain:

\[
f_{gm,k}^{ML} = \frac{2p_{k1,ML}}{p_{0i,ML} + 2p_{k1,ML}}, \tag{20}\]

However, we cannot obtain the 68\% confidence interval of \( f_{gm,k}^{ML} \), defined as \([\sigma_{f_{gm,k}}^{ML}, \sigma_{f_{gm,k}}^{ML}]\), applying the usual error theory, which is based in Gaussianity of variables, because the probability distribution of each \( p_{k1,ML} \) is log-normal. Furthermore, the problem is not analytic and we cannot obtain a mathematical description of the \( f_{gm,k} \) probability distributions. We made Monte Carlo simulations to characterize the probability distribution of each \( f_{gm,k} \). The simulations showed that the \( f_{gm,k} \) distributions can be fit with a log-normal:

\[
P_{LN}(f_{gm,k}) = \frac{\exp\left( -\ln f_{gm,k} - \ln f_{gm,k}^{ML} \right)^2/2\sigma^2}{\sqrt{2\pi f_{gm,k}\sigma}}, \tag{21}\]

where \( \sigma \) is the only free parameter on the fit. Finally, the 68\% confidence interval of \( f_{gm,k}^{ML} \) is given by

\[
\sigma_{f_{gm,k}}^{ML} = e^{-\sigma} f_{gm,k}^{ML}, \tag{22}\]

\[
\sigma_{f_{gm,k}}^{ML} = e^{\sigma} f_{gm,k}^{ML}. \tag{23}\]

3. SIMULATIONS WITH SYNTHETIC CATALOGS

ML methods may suffer from biases in the estimated parameters due to the fact that the maximum likelihood theorem only ensures convergence for a large number of sources. We have used synthetic catalogs in order to test the reliability and accuracy of the ML method. In the following paragraphs, we first explain how we created the synthetic catalogs in a general case, and later define and justify the input parameters used for the synthetic catalogs in this paper. Given the high number of variables used in the following discussion, we provide their precise definitions in Table 1.

We created the synthetic catalogs as follows: first we took \( n \) sources distributed in redshift and asymmetry space following a bidimensional distribution defined by the input probabilities \( p_{kl,\text{in}} \). This process yielded the \( z_{\text{in},i} \) and \( A_{\text{in},i} \) values of the \( n \) sources of our synthetic catalogs, which play the role of \( z_{\text{real},i} \) and \( A_{\text{real},i} \) in equation (4). Next, we applied the experimental errors: following equation (5) we obtained the simulated \( z_{\text{sim},i} \) values, which play the role of \( z_{\text{obs},i} \), as drawn from a Gaussian distribution with mean \( z_{\text{in},i} \) and standard deviation \( \sigma_{z_{\text{in}}} \); the latter plays the role of \( \sigma_{z_{\text{obs}}} \). The value of \( \sigma_{z_{\text{sim}}} \) is a positive value obtained also from a Gaussian distribution with mean \( \overline{\sigma_z} \) and standard deviation \( \sigma_{\sigma_z} \). The process was repeated following equation (6) to obtain the simulated \( A_{\text{sim},i} \) and its standard deviation \( \sigma_{A_{\text{sim}}} \). Finally, we applied the ML method over the synthetic catalog to obtain \( p_{kl,\text{ML}}^{\text{ML}} \) and \( \sigma_{p_{kl,\text{ML}}^{\text{ML}}} \). Summarizing, the input parameters of our simulations were the bidimensional distribution \( p_{kl,\text{in}} \), \( n, z_{\text{in}}, A_{\text{in}}, \sigma_{z_{\text{in}}}, \sigma_{A_{\text{in}}}, \sigma_{z_{\text{obs}}}, \sigma_{A_{\text{obs}}}, \sigma_{A_{\text{in}}}, \sigma_{\sigma_z}, \sigma_{\sigma_A} \), and the output parameters were \( p_{kl,\text{ML}}^{\text{ML}} \) and \( \sigma_{p_{kl,\text{ML}}^{\text{ML}}} \).

We defined three intervals in redshift \((k = 0, 1, 2)\) with \( \Delta z = 0.4 \) and \( z \in [0, 1, 2] \), and two in asymmetry \((l = 0, 1)\) with \( \Delta A = 0.7 \) and \( A \in \{-0.35, 1.05, \} \). Distorted sources with \( A > A_{\text{in}} = 0.35 \) (see C03 for details about the determination of this limit value) are described by \( p_{kl,\text{in}}^{\text{ML}} \) while normal sources by \( p_{kl,0,\text{in}}^{\text{ML}} \). We list in Table 2 the redshift and asymmetry intervals, as well as the probabilities \( p_{kl,\text{in}}^{\text{ML}} \) and \( p_{kl,0,\text{in}}^{\text{ML}} = \ln p_{kl,\text{in}}^{\text{ML}}, \) that define the input bidimensional distribution of our synthetic catalogs. The \( p_{kl,\text{in}}^{\text{ML}} \) values in Table 2 do not match any particular observational determination of these quantities, but they follow the general behavior inferred from observed galaxy merger fractions: highly asymmetric galaxies are less frequent than low-asymmetry galaxies up to \( z = 1.2 \) (Conselice et al. 2003; Cassata et al. 2005; Bridge et al. 2007; Kampczyk et al. 2007), so the \( p_{kl,\text{in}}^{\text{ML}} \) are lower than the \( p_{kl,0,\text{in}}^{\text{ML}} \). The number of highly asymmetric galaxies increases with redshift in the range \( z \in [0, 1.2] \) (Conselice et al. 2003), so \( p_{kl,\text{in}}^{\text{ML}} \) increase with redshift. Several studies present a maximum at intermediate \( z \) in the redshift distribution of galaxies in optically selected samples (e.g., Grazian et al. 2006), so \( p_{kl,\text{in}}^{\text{ML}} + p_{kl,0,\text{in}}^{\text{ML}} \) values have a maximum in the interval \( z = [0.4, 0.8] \). We can check that the \( p_{kl,\text{in}}^{\text{ML}} \) are normalized following equation (10). Although we focused on this particular bidimensional distribution, we carried out the same study with other distributions, and the results were similar.

For convenience we express the experimental dispersions using the dimensionless variables

\[
\sigma_{\text{bin},z} = \frac{\overline{\sigma_z}}{\Delta z}, \tag{24}\]

\[
\sigma_{\text{bin},A} = \frac{\overline{\sigma_A}}{\Delta A}. \tag{25}\]

We used the same value of both variables in each simulation, that is, we used \( \sigma_{\text{bin}} = \sigma_{\text{bin},z} = \sigma_{\text{bin},A} \). Because we fixed the values of \( \Delta z = 0.4 \) and \( \Delta A = 0.7 \), \( \sigma_{\text{bin}} \) unequivocally defines \( \overline{\sigma_z} \) and \( \overline{\sigma_A} \). It is important to notice that, when we work with observational data, the situation is the opposite: our data define \( \overline{\sigma_z} \) and \( \overline{\sigma_A} \), and we should choose the most appropriate values of \( \Delta z \) and \( \Delta A \). We made simulations for \( \sigma_{\text{bin}} = 0 \) as a check corresponding to null experimental errors, \( \sigma_{\text{bin}} = 0.25 \) and 0.5 as
typical observational cases, and $\sigma_{\text{bin}} = 1.0$ as an extreme case to explore the applicability limits of the method. The values of $\sigma_{\text{in}}$ and $\sigma_{\text{class}}$ were a half of $\sigma_z$ and $\sigma_A$, respectively, in all cases.

We ran models with $n = 50, 100$, and $1000$ to check catalog size effects. We took these values because we expect experimental catalogs of a few hundred sources or more and we are interested in the applicability limits of the method to small samples.

In order to study the accuracy of the ML method, we must obtain the distribution followed by the probabilities $p_{\text{ML}}$. Hence, we created several independent synthetic catalogs for each $n$ and $\sigma_{\text{bin}}$ case: we defined a simulation set as $N = 1000$ simulations with each set of input parameters.

The results of the simulations are shown in Figure 1 and in Tables 3, 4, and 5. Figure 1 shows $p_{\text{ML}}$ for samples of $n = 1000$ sources (crosses), with error bars showing their 68% confidence intervals; for comparison, the input probabilities $p_{\text{ML}}$ are shown as black circles, and the $p_{\text{class}}$ obtained by drawing a classical histogram (as defined below in § 3.1), are shown as gray triangles, also for $n = 1000$ catalogs. In Figure 1, panels $a$, $b$, and $c$ correspond to increasing values of the experimental errors, defined in equations (24), (25) and shown in the legend; panels $a$, $b$, and $c$ may be taken to respectively describe “good,” “typical,” and “bad” observational errors as compared to the $z$ and $A$ bin sizes. The top/bottom panels show $p_{\text{ML}}$ for the low/high-asymmetry bins. Within each panel, values for the three redshift bins are shown, as labeled on the horizontal axes. We provide the results in tabular format in Tables 3, 4, and 5.
Before presenting the results of the ML method, we analyze the classical bidimensional histogram of the synthetic catalog without using the ML method. We translate the histogram occupation numbers \( n_{kl} \) to probabilities \( p_{0_{kl};\text{class}} \) using

\[
p_{0_{kl};\text{class}} = \ln \left( \frac{n_{kl}}{\Delta z \Delta A \sum_k \sum_l n_{kl}} \right),
\]

where \( n_{kl} \) is the number of sources in the \([z_k, z_{k+1}) \cup [A_l, A_{l+1})\) bin. The final \( p_{0_{kl};\text{class}} \) probabilities in each simulation set were the median of the \( N \) values of \( p_{0_{kl};\text{class}} \), while standard deviations of \( p_{0_{kl};\text{class}} \), denoted as \( s_{p_{0_{kl};\text{class}}} \), were defined as the dispersion on the \( N \) values of \( p_{0_{kl};\text{class}} \). In Tables 3–5, we can see that the classical bidimensional distribution recovers the input probabilities in the case of null experimental errors and \( n \) large as expected. However, the shape of the input bidimensional distribution begins to deviate when \( \sigma_{\text{bin}} \) increases, as we can also see in Figure 1: the classical bidimensional distribution (gray triangles) is smoothed by experimental errors and does not estimate well the underlying bidimensional distribution (black circles). We study this in detail in § 3.3.

3.2. The ML Method in the Absence of Experimental Errors

We first test that the ML method, in the case of null experimental errors, recovers the input bidimensional distribution, i.e., that it reduces to the classic method. We can see in Tables 3–5 that the values of \( p_{0_{kl};\text{class}} \) and the median of the \( N \) values recovered by the ML method, denoted as \( p_{0_{kl};\text{ML}} \), are the same in all cases. This also happens with the values of \( s_{p_{0_{kl};\text{class}}} \) and the standard deviations of \( p_{0_{kl};\text{ML}} \), denoted as \( s_{p_{0_{kl};\text{ML}}} \). This indicates that the ML method does not introduce systematic effects on the results.

3.3. The ML Method with Non-Null Experimental Errors

We now examine how well the ML and classical methods recover the input probabilities \( p_{0_{kl};\text{in}} \) when non-null experimental errors are included in the synthetic catalogs. We use the \( N = 1000 \) source catalogs as an example, because it is representative of the general trends. The results are shown in Figure 1, and are tabulated in Table 5. It is clear from Figure 1 that \( p_{0_{kl;\text{ML}}}(\text{crosses}) \), recover the input probabilities \( p_{0_{kl;\text{in}}}(\text{black circles}) \) in all cases, including those in which the inserted errors are as large as the bin size (panels c). From Table 5 we see that the values of \( p_{0_{kl;\text{in}}} \) always lie within the 68% confidence interval of the ML method, defined by \( [p_{0_{kl;\text{ML}}}, s_{p_{0_{kl;\text{ML}}} + p_{0_{kl;\text{ML}}} + s_{p_{0_{kl;\text{ML}}}]} \).
This shows that the ML method is reliable. In contrast, the probabilities $p'_{kl,\text{class}}$ derived from the classical histogram (gray triangles in Fig. 1) systematically deviate from the input probabilities. Probabilities are systematically underestimated/overestimated in the low/high-asymmetry bins (upper/lower panels), due to a spillover from the most populated bins (low asymmetries) to the least populated, high-asymmetry bins. Such deviations increase for larger experimental errors. When the errors are as large as the bin size, spillover is so pronounced that the probabilities in the high-asymmetry sample (lower right panel) are nearly equal for the three redshift bins, and all information on the redshift variation of the galaxy merger fractions is lost.

We conclude that the ML method is an unbiased estimator of the input distribution. To put this statement in a more quantitative basis, we carry out a Student’s t-test (Collins 1990 p. 232). We define our estimator as

$$T_{kl,\text{ML}} = \sqrt{N} \frac{p'_{kl,\text{in}} - p'_{kl,\text{ML}}}{s_{p'_{kl,\text{ML}}}}$$

and accept that $p'_{kl,\text{in}} = p'_{kl,\text{ML}}$ with a 99% confidence when $T_{kl,\text{ML}} \leq 2.6$. We define in the same way the variable $T_{kl,\text{class}}$ to study the accuracy of the $p'_{kl,\text{class}}$ as an estimator of the $p'_{kl,\text{in}}$. We calculate the median of the $T_{kl,\text{ML}}$ and $T_{kl,\text{class}}$ for each simulation set, denoted as $T_{\text{ML}}$ and $T_{\text{class}}$, respectively, to make a comparison between different $n$ and $\sigma_{\text{bin}}$.

The results are summarized in Tables 3–5 and in Figure 2. We can see that $T_{\text{ML}}$ is below the confidence level for all $n$ and $\sigma_{\text{bin}}$: the $p'_{kl,\text{ML}}$ are good estimators of the $p'_{kl,\text{in}}$, as wanted. In contrast, the classical method is far from the confidence condition even in the $\sigma_{\text{bin}} = 0.25$ case, and $T_{\text{class}}$ increases with $\sigma_{\text{bin}}$. In addition, having a large $n$ does not improve the results of the classical method: the $p'_{kl,\text{class}}$ values are similar for every $n$, but the errors are reduced when increasing $n$, making $T_{\text{class}}$ higher. That is, having a large observational sample affected by experimental errors does not improve the estimation of $p'_{kl,\text{in}}$, and the $p'_{kl,\text{class}}$ errors are underestimated. This bias affects the galaxy merger fractions obtained from $p'_{kl,\text{class}}$, as we can see on § 3.7.

### 3.4. Study of $\sigma_{kl}$

When we apply the ML method to an observational sample we obtain an estimation of the $p'_{kl,\text{ML}}$ 68% confidence intervals, $[p'_{kl,\text{ML}} - \sigma_{p'_{kl,\text{ML}}}, p'_{kl,\text{ML}} + \sigma_{p'_{kl,\text{ML}}}]$, and we want to know if these
confidence intervals are representative of the \( p_{kl} \) probability distributions. They are representative if the median of the \( N \) values of \( \sigma_{p_{kl}} \), denoted as \( \sigma_{p_{kl,ML}} \), are similar to \( s_{p_{kl,ML}} \). To study this issue we perform a Fisher’s variance test (Collins 1990 p. 234). We define our estimator as

\[
F_{kl} = \frac{\max(\sigma_{p_{kl,ML}}, s_{p_{kl,ML}})^2}{\min(\sigma_{p_{kl,ML}}, s_{p_{kl,ML}})^2}
\]

and accept that \( s_{p_{kl,ML}} = \sigma_{p_{kl,ML}} \) with a 99% of confidence when \( F_{kl} \leq 1.18 \). We calculate the median of the \( F_{kl} \) for each simulation set, denoted as \( F \), to make a comparison between different \( n \) and \( \sigma_{bin} \). The results are summarized in Tables 3–5 and in Figure 3. We can see that \( s_{p_{kl,ML}} = \sigma_{p_{kl,ML}} \) for all \( n \) when \( \sigma_{bin} = 0.25, 0.5 \). Only when \( \sigma_{bin} = 1.0 \) and the samples are small \((n = 50, 100)\) does \( F \) lie above the confidence limits.

These results imply that the ML method supplies reliable confidence intervals of \( p_{kl,ML} \) with thousand-sources samples or, with less sources, if the experimental errors are at most half of the histogram bin size.

The differences between \( s_{p_{kl,ML}} \) and \( \sigma_{p_{kl,ML}} \) have two origins. The main effect comes from the fact that the probability distributions of \( p_{kl,ML} \) are not perfectly Gaussian, and we had assumed Gaussianity to obtain \( \sigma_{p_{kl,ML}} \) analytically. We study this issue in the next section. The other effect is that we evaluated the theoretical values of \( \sigma_{p_{kl,ML}} \) at \( p_{kl,ML} \); the minimization method AMOEBA is not perfect and we may have estimated a local minimum of equation (11) instead the absolute minimum (see § 3.6).

### 3.5. Probability Distributions of \( p_{kl} \)

In the analytical estimation of the \( p_{kl,ML} \) covariances, we assumed that the \( p_{kl,ML} \) probability distributions are Gaussian. To check this assumption we made a histogram of the \( N \) values of \( p_{kl,ML} \) to obtain the shape of the \( p_{kl,ML} \) probability distribution, which we want to approximate by a Gaussian with mean \( \mu_{p_{kl,ML}} \) and standard deviation \( \sigma_{p_{kl,ML}} \). We tested this Gaussian approximation with a Kolmogorov-Smirnov test (Collins 1990 p. 235).

We saw that the Gaussian distribution approximation was valid for all \( \sigma_{bin} \) in the \( n = 1000 \) simulation sets. The situation of the \( n = 50 \) and 100 simulation sets was more complicated. For
n = 100, the $p_{ld,ML}$ Gaussian approximation was valid for all $\sigma_{bin}$, while the $p_{ld,ML}$ started to be non-Gaussian for $\sigma_{bin} = 0.5$, and we could not assume Gaussianity for $\sigma_{bin} = 1.0$. For $n = 50$ simulations we could not assume Gaussian approximation from $\sigma_{bin} = 0.25$ to the $p_{ld,ML}$ and from $\sigma_{bin} = 0.5$ to the $p_{ld,ML}$.

These results emphasize that one must check the Gaussian approximation of the $p_{ld,ML}$ probability distributions in each case. That is, when applying the ML method to an experimental catalog it is essential to make special simulations aimed at verifying the Gaussianity of the recovered probabilities.

### 3.6. The Standard Deviation of the ML Method Due to Iterative Minimization

The iterative minimization method AMOEBA used to obtain the minimum of equation (11) can introduce an error in the determination of $p_{ld,ML}$ if the method converges to a local minimum. In addition, increasing the experimental errors relaxes the conditions over the absolute minimum and makes it more probable that the method converges onto one such local minimum. To study this effect and its importance, we apply the ML method $N = 100$ times over the same catalog, one per simulation set. We define the variable $s_{\tilde{p}_{ld,ML}}$ as the dispersion of the $N$ values of the recovered probabilities $p_{ld,ML}$. We find that the values of $s_{\tilde{p}_{ld,ML}}$ depend on the tolerance and the maximum number of iterations of the minimization method. We take a $10^{-15}$ tolerance and 5000 iterations as optimal values: less tolerance or more iterations does not reduce $s_{\tilde{p}_{ld,ML}}$, but increases the computational time. All final simulations presented in this paper were made with these optimal values. We also find that $s_{\tilde{p}_{ld,ML}}$ increases with $\sigma_{bin}$, but is $\sim$ 5 times smaller than $s_{\tilde{p}_{ld,ML}}$ in the worst experimental error case, so the standard deviations of the probabilities are slightly affected by this effect. Therefore, when applying the ML method to an experimental catalog, it is safe practice to apply it more than once, as a precaution against local solutions and iteration bias.

### 3.7. The Galaxy Merger Fraction

In the previous sections we have seen that the experimental errors modify the input bidimensional distribution, biasing the
classical method estimations, whereas the ML method is able to recover the input bidimensional distribution. In this section we study the general effect and trends that the experimental errors introduce on the galaxy merger fraction determination. To obtain the galaxy merger fraction by the ML method we follow § 2.2. First we determine the galaxy merger fraction $f_{\text{ML}}$ applying equation (16) to the $p_{\text{bl,ML}}$ probabilities in Tables 3–5. Next, we perform Monte Carlo simulations with these $f_{\text{ML}}$ values and the $p_{\text{bl,ML}}$ and $\sigma_{p_{\text{ML}}}$ in Tables 3–5 to characterize the probability distribution of $f_{gm,k}$, obtaining the 68% confidence interval $[\sigma_{f_{gm,k}}, \sigma_{f_{gm,k}}]$ with equations (22) and (23).

The galaxy merger fraction by the classical method is, applying equation (2),

$$f_{\text{class}} = \frac{2e^{\mu_{1,\text{class}}}}{e^{\mu_{1,\text{class}}} + 2e^{\mu_{1,\text{class}}}},$$

while its 68% confidence interval $[f_{gm,k} - \sigma_{f_{\text{class}}}, f_{gm,k} + \sigma_{f_{\text{class}}}]$ is obtained applying the usual error theory to equation (29),

$$\sigma_{f_{\text{class}}} = \frac{2e^{\mu_{1,\text{class}}}}{(e^{\mu_{1,\text{class}}} + 2e^{\mu_{1,\text{class}}})^2} \times \sqrt{s_{p_{1,\text{class}}}^2 + s_{p_{1,\text{class}}}^2}.$$

Because of the experimental error limits of the ML method that we noted in previous sections, we only made this study with the $n = 1000$ simulation sets. We summarize the results in Table 6 and Figure 4. We can see that the classical method gives the worst estimates of the input galaxy merger fraction when the experimental errors increase. We may take as observational reference the $\sigma_{\text{bin}} = 0.25$ case (for example, in Conselice et al. 2003 we have $\sigma_{\text{bin}} \sim 0.2$). In this case, the difference between the input and the classical estimation is $\sim 0.1$ on the first and second redshift intervals, which have the lower input galaxy merger fraction, and $\sim 0.05$ in the third interval. Furthermore, the experimental errors tend to smooth the galaxy merger fraction values. An extreme case is $\sigma_{\text{bin}} = 1$, where the dependency in $z$ has been lost. In addition, the confidence intervals are underestimated and are $\sim 0.05$ in every case. In contrast, the differences between the input and ML method galaxy merger fractions are $\sim 0.01$ in every redshift bin and experimental error case. Furthermore, the 68% confidence intervals are more realistic: in the $\sigma_{\text{bin}} = 0.25, 0.5$ cases they are $\sim 0.05$, while in the $\sigma_{\text{bin}} = 1.0$ case they increase to $\sim 0.1$.

Finally, we also determined the classical galaxy merger fraction in the $n = 50$ and 100 cases, and noticed that the values of $f_{gm,k}$ were similar in each $\sigma_{\text{bin}}$ case: having large samples does

![Figure 2](image2.png)  
**FIG. 2.**—Variation of $T_{\text{ML}}$ (black symbols) and $T_{\text{class}}$ (gray symbols) with dimensionless experimental error size $\sigma_{\text{bin}}$. Triangles are for $n = 50$, circles for $n = 100$, and squares for $n = 1000$ source catalogs. The solid line is the 99% confidence limit $T = 2.6$.

![Figure 3](image3.png)  
**FIG. 3.**—Variation of $F$ with dimensionless experimental error size $\sigma_{\text{bin}}$. Triangles are for $n = 50$, circles for $n = 100$, and squares for $n = 1000$ source catalogs. The solid line is the 99% confidence limit $F = 1.8$. 

### Table 6  
REAL, ML METHOD, AND CLASSIC GALAXY MERGER FRACTION

| $[z_{k-1}, z_{k+1}]$ | $f_{gm,k}^{\text{bin}}$ | $f_{gm,k}^{\text{ML}}$ | $f_{gm,k}^{\text{class}}$ | $\sigma_{\text{bin}} = 0.25$ | $\sigma_{\text{bin}} = 0.5$ | $\sigma_{\text{bin}} = 1.0$ | $\sigma_{\text{bin}} = 0.25$ | $\sigma_{\text{bin}} = 0.5$ | $\sigma_{\text{bin}} = 1.0$ |
|----------------------|------------------------|-----------------------|---------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| [0.0, 0.4)      | 0.3333 | 0.339±0.069 | 0.339±0.048 | 0.358±0.134 | 0.423 ± 0.035 | 0.499 ± 0.038 | 0.575 ± 0.040 |
| [0.4, 0.8)      | 0.3478 | 0.346±0.053 | 0.335±0.040 | 0.343±0.140 | 0.441 ± 0.027 | 0.516 ± 0.029 | 0.583 ± 0.035 |
| [0.8, 1.2]      | 0.4897 | 0.490±0.040 | 0.492±0.030 | 0.480±0.079 | 0.522 ± 0.027 | 0.556 ± 0.030 | 0.595 ± 0.035 |
not improve the results and we may take into account the experimental errors in our analysis to avoid the bias.

4. DETERMINATION OF ANY ONE- OR TWO-DIMENSIONAL DISTRIBUTION BY THE ML METHOD

The method outlined here may easily be applied to the unbiased determination of any bidimensional distribution in the presence of observational errors. For example, the automatic indices $M_{20}$ and $G$ are used in Lotz et al. (2008) to determine the galaxy merger fraction by morphological criteria. We could apply the ML method by defining the variable $MG = G + 0.14M_{20} - 0.33$ and by calling merger systems all sources with $MG > 0$. Similarly, we may apply the ML method to obtain density of sources in color-color diagrams, especially when we have some condition that separates populations, or to determine one-dimensional histogram of any observational magnitude.

For reference, we provide an outline for the application of the ML method to any one- or two-dimensional experimental distribution subject to observational errors:

1. Define the observational catalog. This catalog cannot be restricted to the interval of interest, e.g., $[z_i, z_f]$, because there are sources both with $z_i < z_0$ and $z_j > z_k$ that could belong to a real bidimensional distribution bin within the range of interest due to the observational errors. In general one should include in the sample those sources with $z_i + 2\sigma_i > z_0$ and $z_j - 2\sigma_j < z_k$ to avoid incompleteness effects.

2. Apply the ML method to the observational catalog. First, define the bidimensional distribution bins taking into account the size of the observational errors. Next, minimize equation (11) to obtain the most probable values of $p_{kl}$, $p_{kl,ML}$.

To determine their confidence intervals, calculate the Hessian matrix, equation (13), with the observational data and the previous $p_{kl,ML}$ values. The diagonal elements of the inverse Hessian matrix provide $\sigma_{p_{kl,ML}}$. Notice that we assumed Gaussian experimental errors, equations (5) and (6), in the development of the ML method. If you need to assume other experimental error distributions, you need to recalculate equations (11), (14), and (15) with the new error distributions.

3. Check the results with representative synthetic catalogs. Run simulations with synthetic catalogs to test the accuracy and Gaussianity limits of the method in each particular case following the methodology of §§ 3.3, 3.4, and 3.5. These synthetic catalogs should have the previous $p_{kl,ML}$ as bidimensional distribution input, that is, as $p_{kl,spec}$, and similar characteristics to the experimental ones to fix the other input parameters. For example, synthetic and experimental catalogs should have same number of sources $n_i$ and $\sigma_i$ may be given by the median of the photometric redshift errors in each redshift bin, while, for $\sigma_{z_i}$, one may use the dispersions of these photometric redshift errors. In addition, is important to take into account special cases, e.g., the number of sources with $z_0$, which have $\sigma_z \sim 0$, in each bin, or avoid unphysical values, e.g., negative redshifts.

4. Determine $p_{kl}$, equation (16), and their confidence intervals, equations (18) and (19), in the reliable cases.

5. CONCLUSIONS

We have presented a maximum likelihood method to recover bidimensional distributions of experimental data subject to measurement errors, and applied it to the determination of the galaxy merger fraction based on asymmetry criteria from C03. The Gaussianity of $p_{kl,ML}$ is the strongest condition on the reliability of the method. From the results, taking into account that typical observational catalogs usually have a few hundred sources, and that the probabilities $p_{kl}$ would be small, we conclude that the bin of the bidimensional distribution must be at least twice the typical error in redshift in the observational catalog. Within this quality limit, the ML method can recover with accuracy and reliability the lost information due to the experimental errors. In addition, our results have realistic errors with known shapes, which the classical histograms cannot provide.

The ML method presented here may in principle be extended to as many dimensions as required by the astrophysical problem we are addressing. For instance, if we wish to determine variations in the galaxy merger fraction as a function of galaxy mass, errors in the galaxy mass determination would make objects spill over from one mass bin to the next, biasing the classical histogram approach. The ML method with an added mass axis would solve the problem. Even if we are not seeking to determine the variation of the galaxy merger fraction with mass, our parent sample unavoidably has a boundary (e.g., luminosity, mass, color), and observational errors make objects jump in and out of the sample, hence potentially modifying the shape.
of the distribution we are trying to determine. This extension to higher dimensions is straightforward only when the third variable is independent from the other two. In the case of a third luminosity or mass axis, this is unfortunately not the case: luminosity and mass depend on galaxy redshift, introducing covariances between the variables. Furthermore, luminosity and mass are affected by incompleteness functions, making our problem nonanalytic. We leave the treatment of this problem for future work.

We dedicate this paper to the memory of our six IAC colleagues and friends who met with a fatal accident in Piedra de los Cochinos, Tenerife, in February 2007, with a special thanks to Maurizio Panniello, whose teachings of Python were so important for this paper. This work was supported by the Spanish Programa Nacional de Astronomía y Astrofísica through project AYA2006-12955.

REFERENCES

Arzner, K. et al. 2007, A&A, 468, 501
Bell, E. F., Phleps, S., Somerville, R. S., Wolf, C., Borch, A., & Meisenheimer, K. 2006, ApJ, 652, 270
Bridge, C. R. et al. 2007, ApJ, 659, 931
Bundy, K., Fukugita, M., Ellis, R. S., Kodama, T., & Conselice, C. J. 2004, ApJ, 601, L123
Cassata, P. et al. 2005, MNRAS, 357, 903
Collins, G. W II 1990, Fundamental Numerical Methods and Data Analysis, http://ads.harvard.edu/books/1990fnmd.book/
Conselice, C. J. 2003, ApJS, 147, 1 (C03)
———. 2006, ApJ, 638, 686
Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183
Davidson, R., & Mackinnon, J. 1993, Estimation and inference in econometrics (New York: Oxford University Press)
Davis, M. et al. 2007, ApJ, 660, L1
De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2
De Propris, R., Liske, J., Driver, S. P., Allen, P. D., & Cross, N. J. 2005, AJ, 130, 1516
Efstathiou, G. 2004, MNRAS, 348, 885
García-Dabó, C. E. 2002, Ph.D. Thesis, Universidad Complutense de Madrid
Giavalisco, M. et al. 2004, ApJ, 600, L93
Gottlöber, S., Klypin, A., & Kravtsov, A. V. 2001, ApJ, 546, 223
Governeto, F., Gardner, J. P., Stadel, J., Quinn, T., & Lake, G. 1999, AJ, 117, 1651
Grazian, A. et al. 2006, A&A, 449, 951
Kampczyk, P. et al. 2007, ApJS, 172, 329
Kartaltepe, J. S. et al. 2007, ApJS, 172, 320
Kollat, T. S. et al. 1999, ApJ, 523, L109
Lavery, R. J., Remijan, A., Charmandaris, V., Hayes, R. D., & Ring, A. A. 2004, ApJ, 612, 679
Le Fèvre, O. et al. 2000, MNRAS, 311, 565
Lin, L. et al. 2004, ApJ, 617, L9
Lonsdale, C. J. et al. 2003, PASP, 115, 897
Lotz, J. M. et al. 2008, ApJ, 672, 177
Makarov, D. et al. 2006, AJ 132, 2729
Marsden, J. E., & Tromba, A. J. 1996, Vector Calculus (New York: W.H. Freeman and Company)
Masjedi, M. et al. 2006, ApJ, 644, 54
Mihos, J. C. 1995, ApJ, 438, L75
Naylor, T., & Jeffries, R. D. 2006, MNRAS, 373, 1251
Nelder, J. A., & Mead, R. 1965, Computer Journal, 7(4), 308
Patton, D. R., Carlberg, R. G., Marzke, R. O., Pritchet, C. J., da Costa, L. N., & Pellegrini, P. S. 2000, ApJ, 536, 153
Patton, D. R. et al. 2002, ApJ, 565, 208
Peña, D. 2001, Fundamentos de estadística (Madrid: Alianza Editorial)
Press, W. H. 1995, Numerical Recipies in C, 2nd ed. (New York: Cambridge University Press)
Scoville, N. et al. 2007, ApJS, 172, 1
Sheth, R. K. 2007, MNRAS, 378, 709
Springel, V., Di Matteo, T., & Hernquist, L. 2005, ApJ, 620, L79
Wolf, C., Meisenheimer, K., Rix, H.-W., Borch, A., Dye, S., & Kleinheinrich, M. 2003, A&A, 401, 73