THE ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES, 129:1–31, 2000 July

TESTS OF STATISTICAL METHODS FOR ESTIMATING GALAXY LUMINOSITY FUNCTION AND APPLICATIONS TO THE HUBBLE DEEP FIELD

TSUTOMU T. TAKEUCHI,[1] KOHJI YOSHIKAWA,[1] AND TAKAKO T. ISHII[2]
Department of Astronomy, Faculty of Science, Kyoto University, Sakyo-ku, Kyoto 606–8502, Japan; takeuchi@kusastro.kyoto-u.ac.jp, kohji@kusastro.kyoto-u.ac.jp, ishii@kusastro.kyoto-u.ac.jp
Received 1999 November 28; accepted 2000 February 15

ABSTRACT

We study statistical methods for the estimation of the luminosity function (LF) of galaxies. We focus on four nonparametric estimators: the $1/V_{\max}$ estimator, the stepwise maximum-likelihood estimator, Chołoniewski's estimator, and the improved Lynden-Bell estimator. The performance of the $1/V_{\max}$ estimator has recently been questioned, especially for the faint-end estimation of the LF. We improve these estimators for studies of the distant universe, and examine their performances for various classes of functional forms by Monte Carlo simulations. We also apply these estimation methods to the mock Two-Degree Field (2dF) redshift survey catalog prepared by Cole et al. We find that the $1/V_{\max}$ estimator yields a completely unbiased result if there is no inhomogeneity, but is not robust against clusters or voids. This is consistent with the well-known results, and we do not confirm the bias trend of $1/V_{\max}$ claimed by Willmer in the case of a homogeneous sample. We also find that the other three maximum-likelihood type estimators are quite robust and give consistent results with each other. In practice, we recommend Chołoniewski's estimator for two reasons: (1) it simultaneously provides the shape and normalization of the LF; and (2) it is the fastest among these four estimators, because of its algorithmic simplicity. We then analyze the photometric redshift data of the Hubble Deep Field prepared by Fernández-Soto et al. using the above four methods. We also derive the luminosity density, ρ_L, at the B and I bands. Our B-band estimation is roughly consistent with that of Sawicki, Lin, & Yee, but it is a few times lower at $2.0 < z < 3.0$. The evolution of $\rho_L(I)$ is found to be less prominent.

Subject headings: galaxies: evolution — galaxies: luminosity function, mass function — galaxies: statistics — methods: statistical

1. INTRODUCTION

The luminosity function (LF) of galaxies plays a crucial role for extragalactic astronomy and observational cosmology. It is one of the basic descriptions of the galaxy population itself, and is sometimes treated as a function of color (e.g., Efstathiou, Ellis, & Peterson 1988, hereafter EEP; Metcalfe et al. 1998; Lin et al. 1999) or morphology (e.g., Bingelli, Sandage, & Tammann 1988; Marzke et al. 1998), or other additional parameters of galaxies. It is also essential for interpreting galaxy number counts (e.g., Koo & Kron 1992; Ellis 1997) and for analyzing galaxy clustering (e.g., Strauss & Willick 1995; Efstathiou 1996). Furthermore, the LF is a fundamental test for the theory of galaxy formation (e.g., Baugh, Cole, & Frenk 1996). Recently, the exact shape of the LF has been of particular interest, because it is one of the key issues in the "faint blue galaxy problem" of galaxy number counts (Koo & Kron 1992; Ellis 1997), and may be related to dwarf galaxy formation (e.g., Babul & Rees 1992; Babul & Ferguson 1996; Hogg & Phinney 1997). The evolution of the LF is also important to deriving the cosmic luminosity density, in the context of the "Madau plot," i.e., cosmic star formation density as a function of redshift (e.g., Madau et al. 1996; Cowie, Songaila, & Hu 1996; Sawicki, Lin, & Yee 1997, hereafter SLY97; Pascarelle, Lanzetta, & Fernández-Soto 1998a).

Estimating the galaxy LF from an observational galaxy catalog is fundamental work, but it is not a trivial task. Because of the flux-limited nature of the redshift survey data, the catalogs are inevitably censored, and a suitable statistical technique is required. In the early stages of extragalactic astronomy, the classical estimator, the number of galaxies in a given volume, $\Phi = N/V$, was used to estimate the LF (Hubble 1936). Of course, this is not sufficient for detailed studies, and many experts have proposed ingenious methods. Schmidt (1968) invented the famous $1/V_{\max}$ estimator in the studies of quasar population. Felten (1977) introduced the direction dependence of the magnitude limit. The extension for combining some different catalogs coherently was discussed in Avni & Bahcall (1980). A further extension for examining the evolution of the LF with redshift was proposed by Eales (1993), and the integrated variant of the Eales' estimator was used in the survey of the Hawaii Deep Fields by Cowie et al. (1996). Qin & Xie (1999) also developed this estimator with a similar line of study. The fundamental assumption of this estimator is that the distribution of the objects is spatially uniform. Nowadays this is regarded as a drawback, because we know that the galaxies have strong clustering properties in the large-scale structure. In spite of the drawback, $1/V_{\max}$ estimator has frequently been used for extragalactic studies (e.g., Lilly et al. 1995; Ellis et al. 1996), probably because of its simplicity in calculation.

In order to overcome the difficulty in treating an inhomogeneous galaxy distribution, some density-insensitive methods have been invented. Lynden-Bell (1971) proposed the C^- method, and applied it to the quasar sample of Schmidt (1968). This method is based on a quite sophisticated statistical idea, as we discuss in subsequent sections. Carswell (1973) reported its use in numerical experiments. Jackson (1974) improved the method to combine several different catalogs, and studied the error estimation when the

[1] Research Fellows of the Japan Society for the Promotion of Science.
[2] Kwasan and Hida Observatories, Kyoto University, Yamashina-ku, Kyoto 607–8471, Japan.

LF is expressed as an analytical form. The original method could derive only the shape of the probability density function, but Chołoniewski (1987; hereafter C87) improved the method to obtain the density normalization and trace the density evolution simultaneously. Lynden-Bell himself, and later Felten (1976) and Nicoll & Segal (1983), pointed out the drawback of this method: it cannot work in the faintest regime, where the data points are too sparse. This drawback was basically overcome by the introduction of the smoothing method by Caditz & Petrosian (1993; hereafter CP93). Subba Rao et al. (1996) and Szokoly et al. (1998) used this method in their recent studies of distant galaxies. We note that this method was further generalized by Maloney & Petrosian (1999) to treat the doubly truncated data, but we do not discuss it further here.

The method proposed by Turner (1979) and Kirshner, Oemler, & Schechter (1979) used the ratio of the number of objects in the absolute magnitude interval $[M, M + dM]$ to the number of objects brighter than M, which canceled out the density inhomogeneity. Marinoni et al. (1999) used this method in their analysis of the effect of the local infall motion on the estimation of the LF. Similar estimators were used by Davis & Huchra (1982) and later by de Lapparent, Geller, & Huchra (1989). However, as mentioned in Efstathiou (1996), this estimator does not use the whole sample.

In contrast, some estimation methods using analytical LF models, often called parametric estimation methods, have been developed. Sandage, Tammann, & Yahil (1979, hereafter STY79) introduced the maximum-likelihood method, which was free of the effects induced by density inhomogeneity, in this field by using a parametric Schechter (1976) form for the LF. This parametric form was extended for evolutionary studies of galaxies by Lin et al. (1999). Marshall et al. (1983) presented another parametric estimator that can treat both the LF and the evolution parameter simultaneously, assuming a Poisson distribution of objects on the magnitude-redshift space.

The maximum-likelihood approach was widely used and extended to the methods that did not use analytical forms, often referred to as nonparametric methods. Nicoll & Segal (1983) proposed such a type of estimator and used it for the study of their chronometric cosmology. The estimator, which can be regarded as an advanced version of Nicoll & Segal's method, was invented by Chołoniewski (1986, hereafter C86). This method adopts the same assumption as Marshall et al. (1983), and is regarded as a binned nonparametric version of it. Another stepwise estimator, which was a binned analog of STY79's estimator, was introduced by EEP. This method is called the "stepwise maximum-likelihood method" and now seems to be the most commonly used. Note that not only EEP's but also most of the other estimators are based on the maximum-likelihood principle. EEP's method was also extended to treat density evolution (Heyl et al. 1997; Springel & White 1998).

In spite of the variety of the methods, as we see above, there had only been comparisons of some methods in the literature (e.g., Felten 1976; C86; EEP; Heyl et al. 1997) before the elaborate intercomparison by Willmer (1997, hereafter W97). Statistically detailed discussions are not so frequently seen, either, except in the rigorous work of Petrosian (1992). In W97, each method was examined by Monte Carlo simulations and CfA1 (e.g., de Lapparent et al. 1989) data. The results obtained were fitted by the Schechter form, and W97 discussed the distribution of the estimates by each method after 1000 simulations. Based on the fitting-parameter distributions, W97 reported the bias trends for some estimators. Furthermore, W97 studied the normalization estimates, and concluded that the serious discrepancies between the LFs of local and distant galaxies is not attributable to the differences in the estimators used in the analyses.

Further questions arise after W97. The tests of W97 were restricted to the Schechter-form LF. They considered, for example, the bias in the faint-end slope estimation, and concluded that even for the spatially homogeneous samples, the $1/V_{\max}$ estimator gives biased results. It is often claimed that the faint-end overestimation of the $1/V_{\max}$ estimator is caused by the density inhomogeneity of the Local Supercluster (e.g., Efstathiou 1996). Thus, if galaxies are homogeneously distributed, the estimator is expected to give the correct value. If any subtler problem dwells in the slope estimation, further extensive experiments are required. They also mentioned the binning size selection. For an analysis of the recent very high redshift data, data sparseness should be considered properly.

Recently, the LF of galaxies at extremely high redshift has become available with the aid of large telescope facilities and improved detectors. In addition, redshift surveys have entered upon a new phase by developing the photometric redshift technique. This technique entails much lower observational cost than the spectroscopic survey, and is suitable for analyses of deep photometric data such as the Hubble Deep Field (HDF; Williams et al. 1996). Although some problems are inherent in the technique and in the faint-source finding itself (Ferguson 1998), vast advances have been produced by the method (e.g., Furusawa et al. 2000). The intermediate- to high-redshift results are, however, still controversial with each other.

To settle these problems, reliable and robust analyses of the LF are required. In this paper, we examine and make practical improvements to these estimation methods.[3] Considering the complicated picture of the evolution of galaxy populations, we concentrate our discussions on the nonparametric methods without any assumed functional forms for the LF. We also restrict our concern only to the methods that use the whole sample. We use the mock catalog generated from various shapes of the probability density function (namely, the LF). As we noted above, the density inhomogeneity is a basic property of the galaxy distribution. First we test how accurately these methods reproduce the input density function, by using spatially homogeneous mock catalogs with varying sample size. Next, we examine the estimators by using mock catalogs with a dense cluster and with a large void. We also use the mock Two-Degree Field (2dF) catalog prepared by Cole et al. (1998) in this study. After checking the reliability of each method, we finally apply the methods to the photometric redshift catalog prepared by Fernández-Soto, Lanzetta, & Yahil (1999, hereafter FLY99) and study the evolution of the LF at very large redshift.

This paper is organized as follows. In § 2 we review and discuss the methods and our extensions. Section 3 is devoted to tests of the performance of these methods by

[3] Numerical calculations in this paper are based on the publicly available software package for cosmological study written by one of the authors (K. Y.). The C library can be downloaded from http://www.kusastro.kyoto-u.ac.jp/kohji/research/libcosm/.

mock catalogs. We apply the methods to the photometric redshift catalog and discuss the LF evolution in § 4. Our summary and conclusions are presented in § 5. We briefly introduce the statistical model-selection criterion used in our discussions in the Appendix.

2. NONPARAMETRIC METHODS FOR ESTIMATING THE LUMINOSITY FUNCTION

Before we discuss each method, we define some fundamental quantities. Let M be the absolute magnitude, m the apparent magnitude, and $d_L(z)$ the luminosity distance in Mpc corresponding to redshift z. Then

$$M = m - 5 \log d_L(z) - 25 - K(z) , \tag{1}$$

where $K(z)$ is the K-correction. Here $\log \equiv \log_{10}$. We use the following notation unless otherwise stated: $\phi(M)$ is the luminosity function ($\mathrm{Mpc}^{-3}\ \mathrm{mag}^{-1}$), and N_{obs} is number of detected galaxies in the survey. When we use stepwise estimators, we must select the optimal binning size to suppress the statistical fluctuation (Sturges 1926; Beers 1992 and references therein; Sakamoto, Ishiguro, & Kitagawa 1986; Heyl et al. 1997). We used Akaike's information criteria (AIC; Akaike 1974) in order to select the optimal binning size with the least loss of information (Takeuchi 1999; for a general discussion, see Sakamoto et al. 1986).

2.1. *Schmidt-Eales* ($1/V_{\mathrm{max}}$) *Method*

We discuss here a method used to construct the LF that was originally proposed by Schmidt (1968), well known as the $1/V_{\mathrm{max}}$ method. Eales (1993) developed it further to trace the evolution with redshift. Cowie et al. (1996) used this estimator in an integral form.

We consider the absolute magnitude and redshift range

$$M_l \le M \le M_u ,$$
$$z_l \le z \le z_u , \tag{2}$$

with a survey solid angle Ω, and upper and lower limiting apparent magnitude m_u and m_l, respectively. Then we have

$$\int_{M_l}^{M_u} \phi(M) dM = \sum_{i=1}^{N_{\mathrm{obs}}} \frac{1}{V_{\mathrm{max}}(i)} , \tag{3}$$

$$V_{\mathrm{max}}(i) \equiv \int_{\Omega} \int_{z_{\mathrm{min},i}}^{z_{\mathrm{max},i}} \frac{d^2V}{d\Omega\, dz}\, dz\, d\Omega , \tag{4}$$

where $z_{\mathrm{max},i}$ and $z_{\mathrm{min},i}$ are the upper and lower redshift limits within which a galaxy with the absolute magnitude M_i can be detected in the survey. We note that

$$z_l \le z_{\mathrm{min},i} < z_{\mathrm{max},i} \le z_u .$$

Defining $z(M, m)$ to be the redshift at which a galaxy with absolute magnitude M is observed as an object with apparent magnitude m, we get

$$z_{\mathrm{max},i} = \min \{z_u, z(M_i, m_u)\} , \tag{5}$$

$$z_{\mathrm{min},i} = \max \{z_l, z(M_i, m_l)\} . \tag{6}$$

In fact, both $z_{\mathrm{max},i}$ and $z_{\mathrm{min},i}$ depend on the galaxy spectral energy distributions (SEDs). Thus, we must account for the K-correction when calculating $1/V_{\mathrm{max}}(i)$.

Felten (1976) proved that the Schmidt $1/V_{\mathrm{max}}$ estimator is unbiased, but does not yield a minimum variance. He also proved that the "classical estimator N/V," which is different from the $1/V_{\mathrm{max}}$ estimator, is biased. Willmer (1997) commented that Felten (1976) had shown this estimator to be biased, but it is not exact. A complication of the terminology may have led to this comment.

2.2. *Efstathiou-Ellis-Peterson* (*EEP*) *Method*

In this subsection, we consider the stepwise maximum-likelihood method introduced by EEP. Since the estimator of the EEP method completely cancels the density information, this method requires an independent estimation of the galaxy density. The EEP method uses the form of the LF

$$\phi(M) = \sum_{k=1}^{K} \phi_k W(M_k - M) . \tag{7}$$

The window function, $W(M_\ell - M)$, is defined by

$$W(M_\ell - M) \equiv \begin{cases} 1 & \text{for } M_\ell - \dfrac{\Delta M}{2} \le M \le M_\ell + \dfrac{\Delta M}{2} , \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

According to EEP,[4] the likelihood function is

$$\mathscr{L}(\{\phi_k\}_{k=1,\ldots,K} \mid \{M_i\}_{i=1,\ldots,N_{\mathrm{obs}}}) = \prod_{i=1}^{N_{\mathrm{obs}}} \frac{\sum_{\ell=1}^{K} W(M_\ell - M_i)\phi_\ell}{\sum_{\ell=1}^{K} \phi_\ell H(M_{\mathrm{lim}}(z_i) - M_\ell)\Delta M} , \tag{9}$$

$$H(M_{\mathrm{lim}}(z_i) - M) \equiv \begin{cases} 1 & M_{\mathrm{lim}}(z_i) - \dfrac{\Delta M}{2} > M , \\ \dfrac{M_{\mathrm{lim}}(z_i) - M}{\Delta M} + \dfrac{1}{2} & M_{\mathrm{lim}}(z_i) - \dfrac{\Delta M}{2} \le M < M_{\mathrm{lim}}(z_i) + \dfrac{\Delta M}{2} , \\ 0 & M_{\mathrm{lim}}(z_i) + \dfrac{\Delta M}{2} \le M , \end{cases} \tag{10}$$

where $M_{\mathrm{lim}}(z_i)$ is the absolute magnitude corresponding to the survey limit m_{lim} at redshift z_i.

The logarithmic likelihood is expressed as

$$\ln \mathscr{L} = \sum_{i=1}^{N_{\mathrm{obs}}} \left\{ \sum_{\ell=1}^{K} W(M_\ell - M_i) \ln \phi_\ell - \ln \left[\sum_{\ell=1}^{K} \phi_\ell H(M_{\mathrm{lim}}(z_i) - M_\ell)\Delta M \right] \right\} . \tag{11}$$

Hence, the likelihood equation becomes

$$\frac{\partial \ln \mathscr{L}}{\partial \phi_k} = \sum_{i=1}^{N_{\mathrm{obs}}} \frac{W(M_k - M_i)}{\phi_k} - \sum_{i=1}^{N_{\mathrm{obs}}} \frac{H(M_{\mathrm{lim}}(z_i) - M_k)\Delta M}{\sum_{\ell=1}^{K} \phi_\ell H(M_{\mathrm{lim}}(z_i) - M_\ell)\Delta M} = 0 , \tag{12}$$

which reduces to

$$\phi_k \Delta M = \sum_{i=1}^{N_{\mathrm{obs}}} W(M_k - M_i) \times \left[\sum_{i=1}^{N_{\mathrm{obs}}} \frac{H(M_{\mathrm{lim}}(z_i) - M_k)}{\sum_{\ell=1}^{K} \phi_\ell H(M_{\mathrm{lim}}(z_i) - M_\ell)\Delta M} \right]^{-1} . \tag{13}$$

[4] Koranyi & Strauss (1997) have pointed out that the discreteness of the assumed LF causes a systematic error in the estimation. In order to avoid this effect, we can use the linear extrapolated form. We do not discuss it further in this paper (see the Appendix of Koranyi & Strauss 1997).

This equation can be solved by iteration, and we obtain the maximum-likelihood estimator $\hat{\phi} = \{\hat{\phi}_k\}_{k=1,\ldots,K}$.

As for the normalization of the LF, some estimators have been proposed. We use the following estimator of the mean galaxy density, n, which was used by EEP:

$$n = \frac{1}{V} \sum_{i=1}^{N_{\rm obs}} \frac{1}{\Psi(z_i)} , \tag{14}$$

where V is the maximum volume defined by the largest redshift in the sample, and $\Psi(z)$ is the selection function, defined by

$$\Psi(z) \equiv \frac{\int_{-\infty}^{M_{\rm lim}(z)} \phi(M) dM}{\int_{-\infty}^{\infty} \phi(M) dM} . \tag{15}$$

For further discussion of other normalization estimators, see the Appendix of Davis & Huchra (1982). By combining equations (13) and (14), we get the final results. As Strauss & Willick (1995) pointed out, at large redshift, where the selection function equation (15) is small, the estimator equation (14) becomes noisy. Therefore, in practice a certain cutoff should be introduced in redshift.

2.3. *The Chołoniewski Method*

Here we discuss the method for estimating the LF developed by C86. The advantage of this method is that we can obtain the density and the shape of the LF simultaneously, and can easily examine the galaxy density evolution with redshift. The method explained here is an extended version applicable to the sample with a cosmological scale. In this subsection, we use indices i and j as the labels of the cells.

We again consider the absolute magnitude and redshift range

$$M_l \le M \le M_u ,$$

$$z_l \le z \le z_u ,$$

with a survey solid angle Ω and survey limiting magnitude $m_{\rm lim}$. Let $n(\boldsymbol{r})$ (Mpc^{-3}) be the number density of galaxies in the neighborhood of the position $\boldsymbol{r}$. If we define V as the total comoving volume under consideration, i.e.,

$$V = \int_\Omega \int_{z_l}^{z_u} \frac{d^2V}{d\Omega\, dz}\, dz\, d\Omega , \tag{16}$$

then this leads to the following expression for the mean number density as

$$\bar{n} = \frac{N}{V} , \tag{17}$$

where N is the total number of galaxies within the redshift range $z_l \le z \le z_u$. Here we adopt a statistical model: on the absolute magnitude–position space (M-$\boldsymbol{r}$ space), the galaxy distribution is $f(M, \boldsymbol{r})$, and the probability that we find k galaxies in a volume element $dMdV$ at $(M, \boldsymbol{r})$, P_k, is described as a Poisson distribution,

$$P_k = \frac{e^{-\lambda}\lambda^k}{k!} , \tag{18}$$

$$\lambda = \frac{1}{\bar{n}} f(M, \boldsymbol{r}) dM\, dV . \tag{19}$$

Here

$$\phi(M) = \frac{1}{V} \int_\Omega \int_{z_l}^{z_u} f(M, \boldsymbol{r}) \frac{d^2V}{d\Omega\, dz}\, dz\, d\Omega , \tag{20}$$

$$n(\boldsymbol{r}) = \frac{1}{V} \int_{M_l}^{M_u} f(M, \boldsymbol{r}) dM . \tag{21}$$

If we apply an assumption that the random variables M and $\boldsymbol{r}$ are independent, i.e., $f(M, \boldsymbol{r}) = \psi(M)v(\boldsymbol{r})$, then we obtain

$$\lambda = \frac{1}{\bar{n}} \psi(M) dM\, v(\boldsymbol{r}) dV . \tag{22}$$

We integrate λ over the spherical shell at redshift z and divide the M-z plane into small rectangular cells such that

$$\begin{cases} M_i \le M \le M_{i+1} = M_i + \Delta M & i = 1, \ldots, I , \\ z_j \le z \le z_{j+1} = z_j + \Delta z & j = 1, \ldots, J . \end{cases} \tag{23}$$

Now we see that the problem is to estimate the intensity parameter λ_{ij}, inhomogeneously defined on the M-z plane (see Fig. 1). The probability of finding k_{ij} galaxies in the cell (i, j), $P_{k_{ij}}$, is

$$P_{k_{ij}} = \frac{e^{-\lambda_{ij}} \lambda_{ij}^{k_{ij}}}{k_{ij}!} , \tag{24}$$

which is characterized by the parameter

$$\lambda_{ij} = \int_{M_i}^{M_{i+1}} \int_{V(z_j, z_{j+1})} \lambda \equiv \frac{1}{\bar{n}} \psi_i \Delta M\, v_j V_j , \tag{25}$$

where $V(z_j, z_{j+1})$ is the comoving volume between redshifts z_j and z_{j+1}. Here

$$\psi_i \equiv \frac{1}{\Delta M} \int_{M_i}^{M_{i+1}} \psi(M) dM , \tag{26}$$

$$v_j \equiv \frac{\Omega}{V_j} \int_{z_j}^{z_{j+1}} v(\boldsymbol{r}) \frac{dV}{dz}\, dz , \tag{27}$$

and

$$V_j \equiv \Omega \int_{z_j}^{z_{j+1}} \frac{dV}{dz}\, dz . \tag{28}$$

The likelihood is given by

$$\mathscr{L} = \prod_{(M_i, z_j) \in S} \frac{e^{-\lambda_{ij}} \lambda_{ij}^{k_{ij}}}{k_{ij}!} , \tag{29}$$

and we obtain the log likelihood

$$\ln \mathscr{L} = \sum_{(M_i, z_j) \in S} (k_{ij} \ln \lambda_{ij} - \lambda_{ij} - \ln k_{ij}!) , \tag{30}$$

where S stands for the subset of the M-z plane surrounded by M_u, M_l, z_u, z_l, and the curve $\mathscr{C}$, defined by the selection line

$$M + 5 \log d_L(z) + K(z) + 25 = m_{\rm lim} . \tag{31}$$

We define the following quantities:

$$i_{\max}(j) \equiv \min \{I, i_S(j)\} , \quad j_{\max}(i) \equiv \min \{J, j_S(i)\} , \tag{32}$$

where $M_{i_S(j)} \equiv \{M : \mathscr{C} \cap \{(M, z) : z = z_j\}\}$, and $z_{j_S(i)} \equiv \{z : \mathscr{C} \cap \{(M, z) : M = M_i\}\}$ (see Fig. 1). Using these notations, we

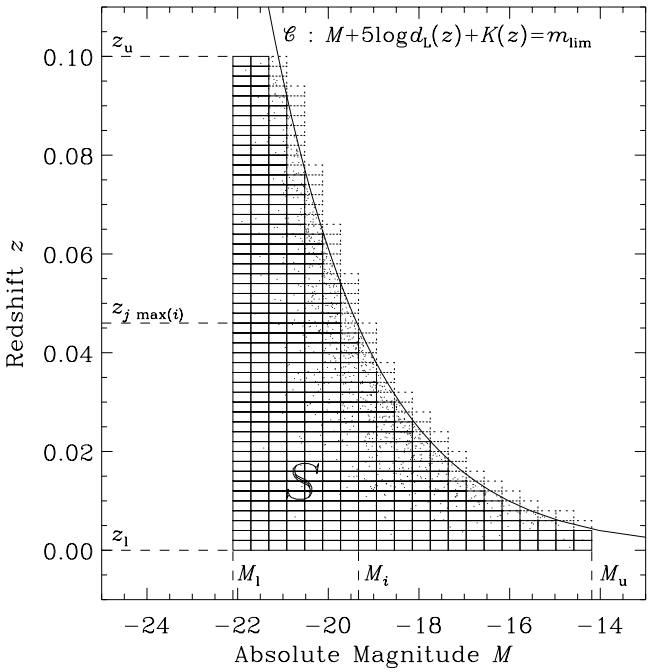

FIG. 1.—Schematic description of the Chołoniewski method

can reduce equation (30) as

$$\ln \mathcal{L} = \sum_{i=1}^{I} \sum_{j=1}^{j_{\max}(i)} (k_{ij} \ln \lambda_{ij} - \lambda_{ij} - \ln k_{ij}!) ,$$
$$= \sum_{j=1}^{J} \sum_{i=1}^{i_{\max}(j)} (k_{ij} \ln \lambda_{ij} - \lambda_{ij} - \ln k_{ij}!) . \quad (33)$$

The maximum-likelihood estimates (MLEs) ($\hat{\psi}_0$, ..., $\hat{\psi}_I$, $\hat{v}_0$, ..., $\hat{v}_J$, $\hat{\bar{n}}$) are the set of solutions that maximizes $\mathcal{L}$. They can be obtained, in practice, by setting the following equations to zero:

$$\frac{\partial \ln \mathcal{L}}{\partial \psi_i} = \sum_{s=1}^{I} \sum_{t=1}^{t_{\max}(s)} \left(k_{st} \frac{1}{\lambda_{st}} \frac{\partial \lambda_{st}}{\partial \psi_i} - \frac{\partial \lambda_{st}}{\partial \psi_i} \right)$$
$$= \sum_{t=1}^{t_{\max}(i)} \left(\frac{k_{it}}{\psi_i} - \frac{1}{\bar{n}} \Delta M\, v_t\, V_t \right) = 0 , \quad (34)$$

$$\frac{\partial \ln \mathcal{L}}{\partial v_j} = \sum_{t=1}^{J} \sum_{s=1}^{s_{\max}(t)} \left(k_{st} \frac{1}{\lambda_{st}} \frac{\partial \lambda_{st}}{\partial v_j} - \frac{\partial \lambda_{st}}{\partial v_j} \right)$$
$$= \sum_{s=1}^{s_{\max}(j)} \left(\frac{k_{sj}}{v_j} - \frac{1}{\bar{n}} \psi_s \Delta M V_j \right) = 0 . \quad (35)$$

Thus we have a set of equations that are referred to as likelihood equations:

$$\frac{\Delta M}{\bar{n}} \psi_i \sum_{t=1}^{t_{\max}(i)} v_t V_t = \sum_{t=1}^{t_{\max}(i)} k_{it} , \quad (36)$$

$$\frac{\Delta M}{\bar{n}} v_j V_j \sum_{s=1}^{s_{\max}(j)} \psi_s = \sum_{s=1}^{s_{\max}(j)} k_{sj} . \quad (37)$$

These equations are solved by iterative procedure. At this stage, these solutions obtained here are not exactly the MLEs themselves, but relative values. We need one more step to obtain absolute values. We denote the relative solutions by a tilde (~) and exact MLEs by a circumflex (^). From equations (36) and (37), we have

$$\tilde{\psi}_i = \frac{\bar{n}}{\Delta M} \frac{\sum_{t=1}^{t_{\max}(i)} k_{it}}{\sum_{t=1}^{t_{\max}(i)} \tilde{v}_t V_t}$$
$$= \sum_{t=1}^{t_{\max}(i)} k_{it} \left[\sum_{t=1}^{t_{\max}(i)} \frac{\sum_{s=1}^{s_{\max}(t)} k_{st}}{\sum_{s=1}^{s_{\max}(t)} \tilde{\psi}_s} \right]^{-1} , \quad (38)$$

$$\tilde{v}_j V_j = \frac{\bar{n}}{\Delta M} \frac{\sum_{s=1}^{s_{\max}(j)} k_{sj}}{\sum_{s=1}^{s_{\max}(j)} \tilde{\psi}_s}$$
$$= \sum_{s=1}^{s_{\max}(j)} k_{sj} \left[\sum_{s=1}^{s_{\max}(j)} \frac{\sum_{t=1}^{t_{\max}(s)} k_{st}}{\sum_{t=1}^{t_{\max}(s)} \tilde{v}_t V_t} \right]^{-1} . \quad (39)$$

We then properly normalize these solutions. Clearly, it follows that

$$\sum_{(M_i, z_j) \in S} \lambda_{ij} = \Delta M \sum_{(M_i, z_j) \in S} \hat{\psi}_i \hat{v}_j V_j = N_{\rm obs} . \quad (40)$$

If we set $\tilde{\psi}_i \tilde{v}_j = w \hat{\psi}_i \hat{v}_j$, then we straightforwardly obtain the numerical factor w by equation (40):

$$w = \frac{N_{\rm obs}}{\Delta M \sum_{(M_i, z_j) \in S} (\tilde{\psi}_i \tilde{v}_j V_j)} . \quad (41)$$

Now we obtain the LF $\phi(M)$ and density $n(z)$ as

$$\phi(M_i) = \frac{1}{V} \hat{\psi}_i \sum_{j=1}^{J} \hat{v}_j V_j , \quad (42)$$

$$n(z_i) = \frac{\Delta M}{V} \hat{v}_j V_j \sum_{i=1}^{I} \hat{\psi}_i . \quad (43)$$

2.4. *The Lynden-Bell–Chołoniewski–Caditz–Petrosian (LCCP) Method*

In this section, we discuss the method originally introduced by Lynden-Bell (1971) as the "C^- method." The estimator of this method is an analog of the Kaplan-Meier estimator used for censored data analyses, such as survival analysis (e.g., Feigelson & Nelson 1985; Feigelson 1992; Babu & Feigelson 1996; for a discussion of survival analysis itself, see, e.g., Kleinbaum 1996). This method may be the most natural application of the nonparametric statistics to the problem (e.g., Petrosian 1992). The version of the method rederived by C87 was improved so that it could estimate the LF and density evolution of galaxies at the same time. In addition, the derivation of the estimator was much simplified. The original method was invented to estimate the cumulative LF as a step function; thus, the differential LF was described as a weighted sum of Dirac's δ-function. However, this form is obviously not practical, and C87 suggested smoothing the LF. In modern statistics, the kernel estimator is used in the problem of nonparametric density estimation (Silverman 1986; Lehmann 1999). The kernel is a smooth function that is used as a substitute for the δ function, in order to keep the estimated density function smooth. This improvement was introduced to the LF estimation problem by CP93, and used for a photometric redshift catalog by Subba Rao et al. (1996).[5]

We unify these improvements, and here show the practically convenient calculation, which we call the "LCCP method" after the names of the above contributors. We use the same notations for luminosity function, galaxy number

[5] However, we note that Subba Rao et al.'s equation (6) erroneously includes an extra exponential.

density, distribution of galaxies, etc., and we consider the same absolute magnitude and redshift ranges as in § 2.3. However, we note that in this subsection, indices represent the labels of galaxies. This method is completely free of binning procedure.

For the following discussion, we suppose that the galaxies are ordered as $M_k \leq M_{k+1}$. In the LCCP method, the independence assumption is also adopted for M and z, which leads to the expression

$$f(M, z) = \psi(M)v(z) .$$

The empirical distribution (distribution of observational data) is expressed as

$$f_{\rm obs}(M, z) = \sum_{k=1}^{N_{\rm obs}} \delta(M - M_k, z - z_k) , \tag{44}$$

where again $N_{\rm obs}$ is the observed sample size, and let

$$\psi(M) = \sum_{i=1}^{N_{\rm obs}} \psi_i \, \delta(M - M_i) , \tag{45}$$

$$v(z) = \sum_{j=1}^{N_{\rm obs}} v_j \, \delta(z - z_j) . \tag{46}$$

Then the empirical distribution is

$$\begin{aligned} f_{\rm obs}(M, z) &= \sum_{i=1}^{N_{\rm obs}} \psi_i \, \delta(M - M_i) \sum_{j=1}^{N_{\rm obs}} v_j \, \delta(z - z_j) \chi_S(M, z) \\ &= \sum_{(i, j) \in S} \psi_i \, v_j \, \delta(M - M_i) \delta(z - z_j) . \end{aligned} \tag{47}$$

Here, $\chi_S(M, z)$ is the characteristic function of the set S, defined as

$$\chi_S(M, z) \equiv \begin{cases} 0 & (M, z) \notin S , \\ 1 & (M, z) \in S . \end{cases}$$

In the following discussions, the quantities $M_{\rm max}(j)$ and $z_{\rm max}(i)$ are defined as

$$\begin{aligned} M_{\rm max}(j) &\equiv \min \{M_u, M_{S(j)}\} , \\ z_{\rm max}(i) &\equiv \min \{z_u, z_{S(i)}\} , \end{aligned} \tag{48}$$

where $M_{S(j)} \equiv \{M : \mathscr{C} \cap \{(M, z) : z = z_j\}\}$, and $z_{S(i)} \equiv \{z : \mathscr{C} \cap \{(M, z) : M = M_i\}\}$. Although they look like those used in § 2.3, we note again that the indices are of galaxies. These are schematically described in Figure 2. Integration of equation (47) over the interval $[M_l, M_u]$, $[z_k - \epsilon, z_k + \epsilon]$ $(\epsilon > 0)$ gives

$$\begin{aligned} &\int_{M_l}^{M_u} \int_{z_k - \epsilon}^{z_k + \epsilon} \sum_{\ell=1}^{N_{\rm obs}} \delta(M - M_\ell, z - z_\ell) dz \, dM = 1 \\ &= \int_{M_l}^{M_u} \int_{z_k - \epsilon}^{z_k + \epsilon} \sum_{(i, j) \in S} \psi_i \, v_j \, \delta(M - M_i) \delta(z - z_j) dz \, dM \\ &= v_k \sum_{i=1}^{M_i < M_{\rm max}(k)} \psi_i \; ; \end{aligned}$$

therefore

$$v_j \sum_{i=1}^{M_i < M_{\rm max}(j)} \psi_i = 1 , \quad j = 1, \ldots, N_{\rm obs} . \tag{49}$$

Similarly, integration over $[M_k - \epsilon, M_k + \epsilon]$, $[z_l, z_u]$ gives

$$\psi_i \sum_{j=1}^{z_j < z_{\rm max}(i)} v_j = 1 , \quad i = 1, \ldots, N_{\rm obs} . \tag{50}$$

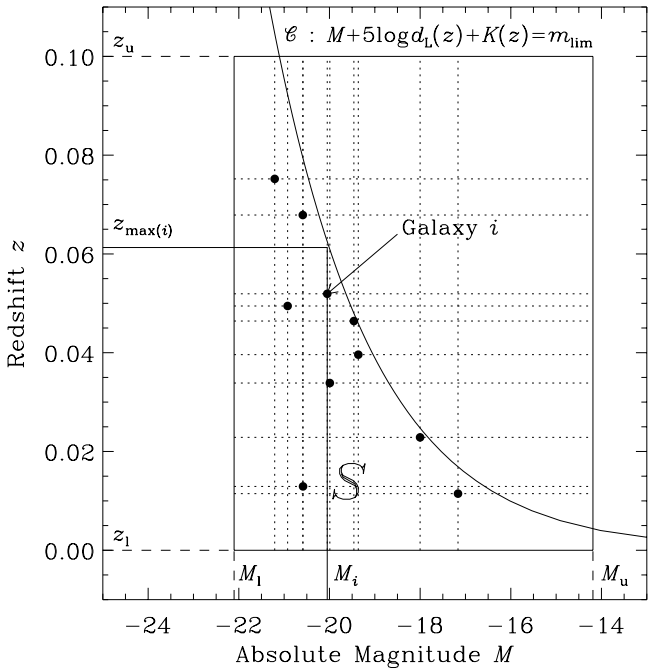

FIG. 2.—Schematic description of the Lynden-Bell–Chołoniewski–Caditz–Petrosian method.

Formally, we can obtain $\{\psi_i\}_{i=1,\ldots,N_{\rm obs}}$ and $\{v_j\}_{j=1,\ldots,N_{\rm obs}}$ by solving the equations (49) and (50), and the estimates of the real galaxy distribution, $f(M, z)$, as

$$f(M, z) = \sum_{i=1}^{N_{\rm obs}} \sum_{j=1}^{N_{\rm obs}} \psi_i \, \delta(M - M_i) v_j \, \delta(z - z_j) . \tag{51}$$

Thus,

$$\phi(M) = \frac{1}{V} \int_{z_l}^{z_u} f(M, z) dz = \frac{1}{V} \sum_{i=1}^{N_{\rm obs}} \psi_i \; \delta(M - M_i) \sum_{j=1}^{N_{\rm obs}} v_j , \tag{52}$$

$$n(z) = \frac{dz}{dV} \int_{M_l}^{M_u} f(M, z) dM = \frac{dz}{dV} \sum_{j=1}^{N_{\rm obs}} v_j \, \delta(z - z_j) \sum_{i=1}^{N_{\rm obs}} \psi_i , \tag{53}$$

where V is the volume considered. The total number of galaxies, N, is

$$N = \sum_{i=1}^{N_{\rm obs}} \psi_i \sum_{j=1}^{N_{\rm obs}} v_j . \tag{54}$$

These solutions are MLEs, as discussed in C87.

In spite of the clarity of the derivation, it is in fact not an easy task to solve equations (49) and (50) numerically if the data set size, $N_{\rm obs}$, is large. Thus, we use the usual C estimators together, in order to calculate the Chołoniewski coefficients more easily. The Lynden-Bell C^- function, $C^-(M_k)$, is the number of galaxies in the region

$$\begin{aligned} &M_l \leq M < M_k , \\ &z_l \leq z \leq z_{\rm max}(k) . \end{aligned} \tag{55}$$

Let

$$C_k \equiv C^-(M_k) , \quad k = 1, \ldots, N_{\rm obs} . \tag{56}$$

Then, using equations (49) and (50), we have

$$C_k + 1 = \sum_{i=1}^{k} \psi_i \sum_{j=1}^{z_j < z_{\max}(k)} v_j = \sum_{i=1}^{k} \frac{\psi_i}{\psi_k}, \tag{57}$$

and we obtain the following recursion relation:

$$\psi_{k+1} = \frac{C_k + 1}{C_{k+1}} \psi_k \ . \tag{58}$$

Thus, the distribution function of M (cumulative LF), $\Phi(M)$, is

$$\Phi(M) \propto \sum_{k=1}^{M_k < M} \psi_k = \psi_1 \prod_{k=1}^{M_k < M} \frac{C_k + 1}{C_k} \ . \tag{59}$$

In the real procedure, we set $(C_1 + 1)/C_1 = 1$, so the product in the above equation begins with $k = 2$. We can prove the second step of the above equation by mathematical induction. This is equivalent to Lynden-Bell's solution (C87). We obtain the weight, $\{\psi_i\}_{i=1,\ldots,N_{\rm obs}}$, by equation (58), and we are able to calculate the density weight, $\{v_j\}_{j=1,\ldots,N_{\rm obs}}$, by equation (50).

As we mentioned above, the weighted sum of the δ-function is not a practically useful form, and random fluctuation would be serious in the region where the data points are sparse. Therefore, the kernel estimator, which is often used in modern nonparametric density estimation, was introduced by CP93. This estimator is simply obtained by replacing the δ-function with a smooth kernel function, κ, as

$$f(M, z) = \sum_{i=1}^{N_{\rm obs}} \sum_{j=1}^{N_{\rm obs}} \psi_i v_j \frac{1}{h_M h_z} \kappa\left(\frac{M - M_i}{h_M}\right) \kappa\left(\frac{z - z_j}{h_z}\right) . \tag{60}$$

The minimum value of the "smoothing scale" h is restricted by the observational uncertainty, which was used by Subba Rao et al. (1996), but it does not provide sufficient smoothing in general (CP93). The optimal value of h_M or h_z can be estimated as

$$h_M \sim \max \{M_{i+1} - M_i\}_{i=1,\ldots,N_{\rm obs}} ,$$
$$h_z \sim \max \{z_{j+1} - z_j\}_{j=1,\ldots,N_{\rm obs}} . \tag{61}$$

It is obvious that the larger the data size $N_{\rm obs}$ is, the smaller the smoothing scale becomes. Furthermore, CP93 discussed the effect of the kernel shape on the estimates. Now it is known that the best shape for the kernel is parabolic, the so-called Epanechnikov kernel (Epanechnikov 1969), because it gives the minimum variance (Lehmann 1999; van Es 1991):

$$\kappa(x) = \tfrac{3}{4}(1 - x^2) . \tag{62}$$

It should be noted that, in principle, the kernel estimator is asymptotically biased, i.e., the expectation value is slightly different from the true value even if the sample size is large.

3. TEST OF THE METHODS BY SIMULATION

3.1. *Numerical Examination with Mock Catalogs*

The validity of estimation methods for the LF is often examined by mathematical statistics. For example, their statistical unbiasedness and statistical convergence were discussed in many early works (e.g., Felten 1976). However, quantitative evaluation frequently appears to be difficult by such an approach, and numerical examination is quite important. Jackson (1974) used numerical experiments, as well as the analytical error estimation from Fisher's information matrix (see Stuart, Ord, & Arnold 1999), in the study of the quasar LF, and EEP also checked the errors of their method by Monte Carlo simulations as well as by a traditional information matrix approach. Mobasher, Sharples, & Ellis (1993) performed Monte Carlo error estimation to test the special method developed to construct the LF at a certain waveband from the data selected at another wavelength. Heyl et al. (1997) examined the effect of galaxy clustering on the LF estimation by their extended EEP method. However, a computer-aided extensive comparison between the estimators had not been performed until the work of W97. They discussed the performance of several estimators when the LF is represented by the Schechter form

$$\phi(M)dM = 0.4 \ln 10 \, \phi_* \, 10^{-0.4(\alpha+1)(M-M_*)} \times \exp\left[-10^{-0.4(M-M_*)}\right] dM , \tag{63}$$

and tested the results in some cases with different Schechter parameters. Their main conclusions are as follows:

1. The STY79 and C^- methods are the best.
2. The $1/V_{\rm max}$ method gives biased results and tends to give higher values for the faint-end slope *even for spatially homogeneous samples.*
3. The STY79 fit tends to underestimate the faint-end slope.
4. The mean densities (normalization of the LF) recovered by most estimators are lower than the input values by factors of up to 20%.

Among these, the second finding looks most strange, because as we mentioned in § 2.1, Felten has proved mathematically that the $1/V_{\rm max}$ estimator is unbiased when the homogeneous assumption holds. The $1/V_{\rm max}$ method is quite frequently used in the estimation of the LFs of quasars, clusters of galaxies, etc., and if W97's claim is true, some widely accepted conclusions must be significantly changed. Thus, it is necessary to examine the estimators further, not only for the Schechter form but for various shapes of the LF, in order to clarify the trends of the results.

In this section, we test the four estimators discussed in the previous section by using simulated mock galaxy samples with a variety of LFs that have the following functional forms:

A. Uniform distribution,
B. Power-law form that increases toward fainter magnitude.
C. Power-law form that decreases toward fainter magnitude.
D. Gaussian distribution (with standard deviation 1.67 mag).
E. Schechter form (flat faint-end slope: $\alpha = -1.1$).
F. Schechter form (steep faint-end slope: $\alpha = -1.6$).

All these forms have a magnitude range $M = [-24, -14]$ (Fig. 3). The first three forms are designed to examine the effect of LF slope on estimation, and form D, the Gaussian, is to check the effect of the curvature of the function. The power-law LF of form B appears ubiquitously in various types of objects. Form C appears unrealistic, but we added it in order to make a thorough investigation. Form D is interesting because an approximate Gaussian form is often

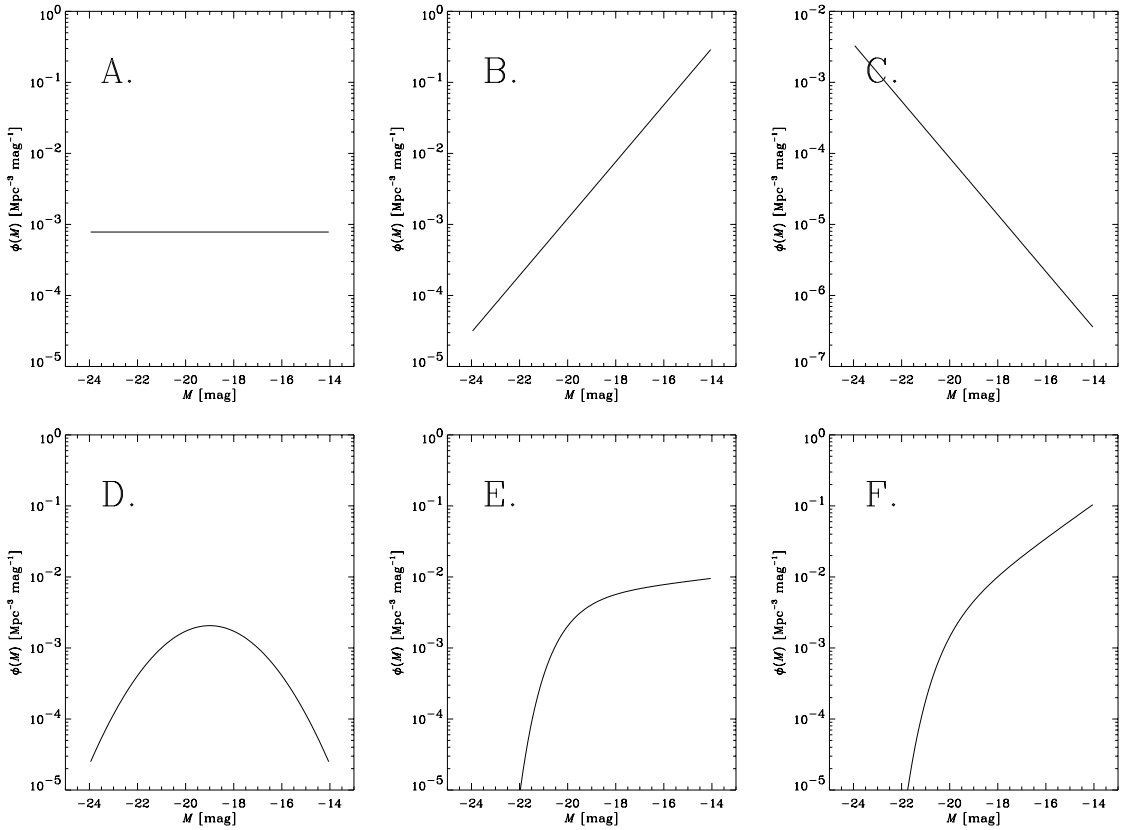

FIG. 3.—Simulated functional forms for the luminosity functions. The density scale is arbitrary because it depends on the adopted sample size, and we set the scale for the case of size ~ 1000.

found in the LFs of individual galaxy types. We applied the Box-Muller method (Box & Muller 1958) to generate a Gaussian distribution from a uniform random number, and von Neumann's acceptance-rejection method to obtain other distributions (see Knuth 1998 for details). We set the sample sizes at ~ 100 and ~ 1000 in the case of the homogeneous sample, to study the behavior of the statistical estimators with galaxy number. Here "sample size" means the detected number of galaxies after magnitude-selection (observation) procedures. Therefore, the underlying population density is different for each LF form. The estimation of galaxy spatial density is an important part of the derivation of the LF. To be estimated is the total galaxy number, including the galaxies too faint to be observed. In our simulations, we stochastically produce galaxies according to the assumed LF, distribute them in space, calculate their observed flux, and judge whether or not they could be observed. Therefore, the total number corresponds to the number of Monte Carlo trials. We fixed the number of trials through one sequence of simulations with a certain LF shape and spatial density.

3.1.1. *Mock Catalog with Spatially Homogeneous Distribution*

First we construct a set of mock galaxy samples with a spatially homogeneous distribution in order to investigate the bias trend of the estimators, especially for the Schmidt and Eales $1/V_{\max}$. We set the redshift range up to 0.1, and we adopted the Hubble parameter $H_0 = 75\ \mathrm{s}^{-1}\ \mathrm{Mpc}^{-1}$, $\Omega_0 = 0.2$ ($q_0 = 0.1$), $\lambda_0 = 0$, and limiting magnitude $m_{\lim} = 13$ mag in the series of simulations. No K-correction is considered here. We constructed 100 representations for each LF form and sample size, and applied the four estimators to each sample.

3.1.2. *Mock Catalog with a Dense Cluster and with a Void*

We next investigate the response against the density inhomogeneity of galaxies. We consider some extreme cases for clearer understanding. For the case with density enhancement, we constructed a series of mock catalogs with a dense spherical clump, to which half of the galaxies belong. The clump lies at a distance of 0.8 Mpc, and its radius is 0.8 Mpc. We call the mock catalog the "cluster sample." An example of the spatial configuration of galaxies of a cluster sample is described in Figure 4. We then also constructed a set of the mock catalogs with a large spherical void without galaxies. The void lies at a distance of 0.8 Mpc, and its radius is 1.6 Mpc. We call this mock catalog the "void sample." The overall underlying density of cluster and void samples defined in a considered volume is the same as the homogeneous samples for each LF shape, i.e., we set the number of Monte Carlo trials to be the same as that of the homogeneous sample for each LF shape. Therefore, the observed sample size of the cluster sample is larger than that of the homogeneous sample, because we put the dense clump in the considered volume. In the case of the void sample, the observed galaxy number is smaller than that of the homogeneous sample.

3.1.3. *Results*

The results for the 1000 galaxy samples are shown in Figures 5–10. The solid lines represent the input distributions, and the symbols show the averages of the estimates.

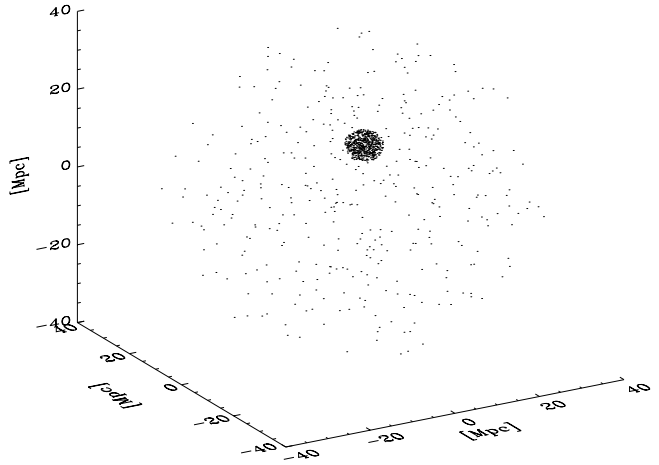

FIG. 4.—Example of the spatial configuration of galaxies of a cluster sample.

The error bars depict the standard deviations of the mean of the estimates for 100 representations. Figures 5*a*, 5*b*, and 5*c* show the results from the spatially homogeneous sample, from the cluster sample, and from the void sample, respectively. This is the same in Figures 6–10.

At a glance, we see that all estimators give results that are consistent with each other, and we do not find any bias trends in our numerical experiments for any LF forms in the case of homogeneous samples. For cluster samples, the $1/V_{\max}$ method yields strongly distorted estimations, as widely recognized. The overestimation corresponding to the dense clump clearly appeared in the $1/V_{\max}$ results. In contrast, the other three estimators were not affected by the dense cluster at all. The estimates appeared to be consistent with each other, and showed perfect agreement with the input LFs. The $1/V_{\max}$ method was also affected by the large void, and gave underestimated results.

Large fluctuations appear at the faint end of the LF, because the number of available data points is small, especially in the case of the LF forms C and D. We can obtain statistically stable estimates if the slope is properly steep, and the shallower the slope is, the larger the fluctuation becomes. This is clearly shown in Figures 5–10.

In principle, the error bar of the Chołoniewski method is larger than those of the other methods, because the method subdivides the M-z plane in both M and z. This procedure enables us to estimate the shape, the normalization, and the evolution of the LF at the same time. On the other hand, this becomes a drawback when the data set is small, because the shot noise dominates. Therefore, we cannot expect a firm estimation with the Chołoniewski method when the sample size is smaller than 100.

Here we mention the calculation time that each method needs for the same sample size. Because of its algorithmic simplicity, the Chołoniewski method is the fastest among the four methods. When we analyze the 1000 mock data, the relative calculation times of the $1/V_{\max}$, EEP, and LCCP methods normalized with that of the Chołoniewski method are 2.76, 2.73, and 1.87, respectively. This advantage is quite significant when we treat a large sample of $\sim 10^4$–10^5 galaxies. We estimate the LFs from large data sets of sample size 250,000 in § 3.2 by the $1/V_{\max}$, EEP, and Chołoniewski methods. The relative calculation times for the $1/V_{\max}$ and EEP methods normalized with that of the Chołoniewski method are, in this case, 15.01 and 131.74, respectively. The EEP method takes a longer calculation time because it needs more iterations in the procedure than the others. The $1/V_{\max}$ method derives the maximum volume, $V_{\max}$, for each galaxy, and also needs some calculation time. The LCCP method requires a large stack for the data-sorting procedure, which is a requirement of this method. Thus, we stress that the Chołoniewski method is the most economical from the standpoint of practical computing.

Figures 11–16 are the same as Figures 5–10, except that the data size is 100. We see that it is often not possible to determine the faint end of the LF accurately for such small data sets. The fluctuation becomes larger than the result of the 1000 galaxy sample, but we did not find the systematic bias trend from our results. Thus, we conclude that when the galaxy distribution is homogeneous, all four estimators provide consistent and correct results, even the $1/V_{\max}$ estimator.

3.2. *Mock 2dF Redshift Catalog*

The Anglo–Australian Two-Degree Field (2dF) galaxy redshift survey is now underway.[6] This survey will measure 250,000 redshifts, up to $z \sim 0.2$, and will be complete to an extinction-corrected apparent magnitude of $b_J < 19.45$ mag. In order to develop statistical methods and faster algorithms for the analyses of such large upcoming redshift surveys, Cole et al. (1998) prepared an extensive set of mock 2dF catalogs constructed from a series of large cosmological N-body simulations. The simulations span a wide range of cosmological models, with various values of the density parameters, Ω_0, the cosmological constant, λ_0, and the shape parameter, Γ, and amplitude of the density fluctuation, σ_8. The LF is assumed to be a Schechter form with the parameters reported by APM-Stromlo bright galaxy survey (Loveday et al. 1992), $M_{b_J^*} - 5 \log h = -19.5$ mag, $\alpha = -0.97$, and $\phi_* = 1.4 \times 10^{-2}\ h^3\ \mathrm{Mpc}^{-3}$. The K-correction is assumed to be canceled by evolutionary correction.

We applied the three methods to the mock 2dF catalog in order to see how accurately they can reproduce the true LF when they are used in the analysis of realistic large-redshift surveys. We did not use the LCCP method for this sample. When we treat such a large catalog, the advantage of the Chołoniewski methods is extremely significant. We also focused on the difference between the real-space data and the redshift-space data, which is affected by the redshift distortion. The redshift distortion causes a scatter in the estimated luminosities of galaxies. In this study, we used three mock catalogs, labeled E1 (Einstein–de Sitter: $\Omega_0 = 1$, $\lambda_0 = 0$, $\Gamma = 0.5$, $\sigma_8 = 0.55$), L3S ($\Omega_0 = 0.3$, $\lambda_0 = 0.7$, $\Gamma = 0.25$, $\sigma_8 = 1.13$), and O3S ($\Omega_0 = 0.3$, $\lambda_0 = 0$, $\Gamma = 0.25$, $\sigma_8 = 1.13$). The catalogs we selected are all cluster-normalized, i.e., the amplitude of the initial power spectrum is set to reproduce the present abundance of rich galaxy clusters in the local universe (e.g., Viana & Liddle 1996; Kitayama & Suto 1997) and $h = \Gamma/\Omega_0$.

We compare the input LF and the estimated LF in Figures 17, 18, and 19. Figure 17 shows the LF derived from the Einstein–de Sitter (EdS) universe, Figure 18 shows the LF derived from the L3S data, and Figure 19 shows the LF derived from the O3S data. The left panels in these figures

[6] See http://msowww.anu.edu.au/xxxcolless/2dF for recent status.

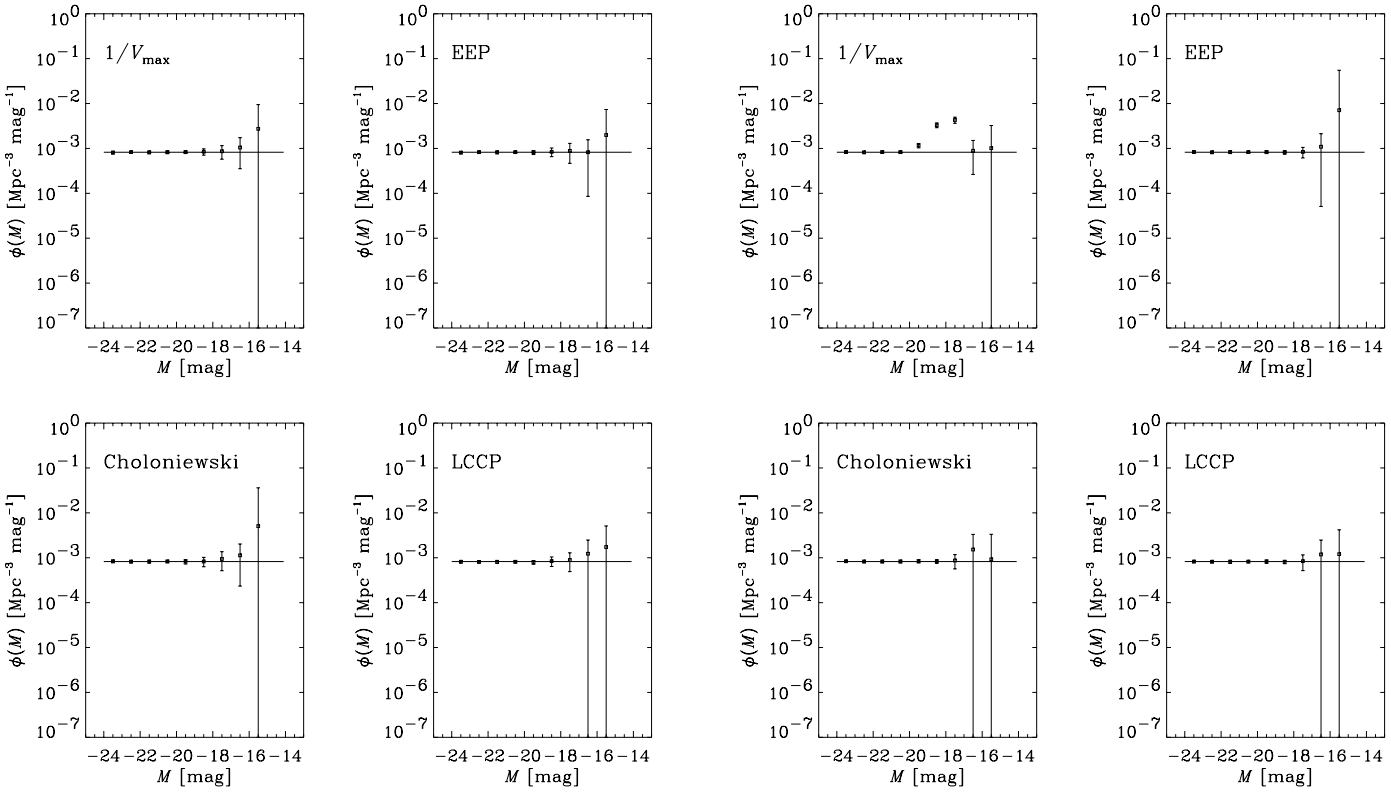

FIG. 5*a*

FIG. 5*b*

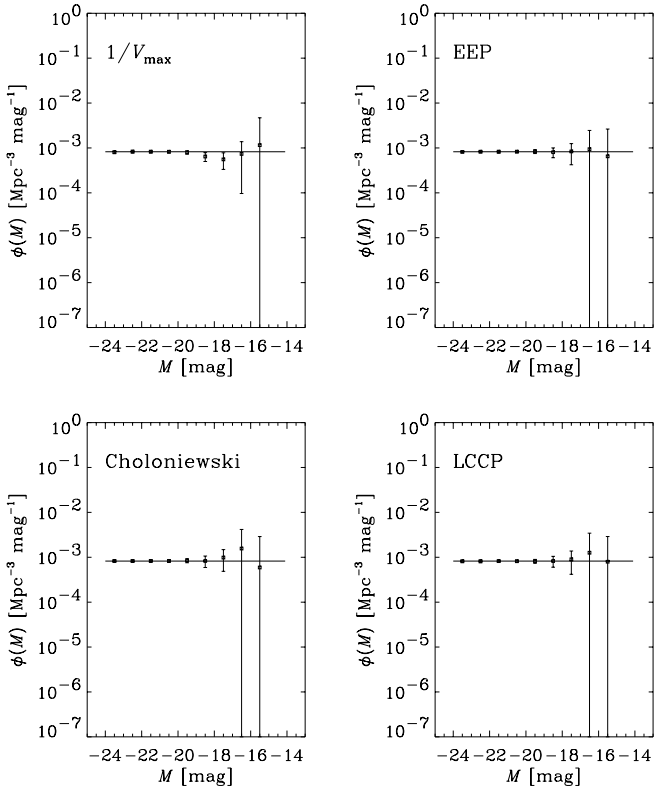

FIG. 5*c*

FIG. 5.—Estimates of the luminosity function (LF) for mock samples with size $\sim$1000, shown for (*a*) the LF of the spatially homogeneous sample; (*b*) the LF estimated from the cluster sample; and (*c*) the LF estimated from the void sample. The input luminosity function has the functional form A.

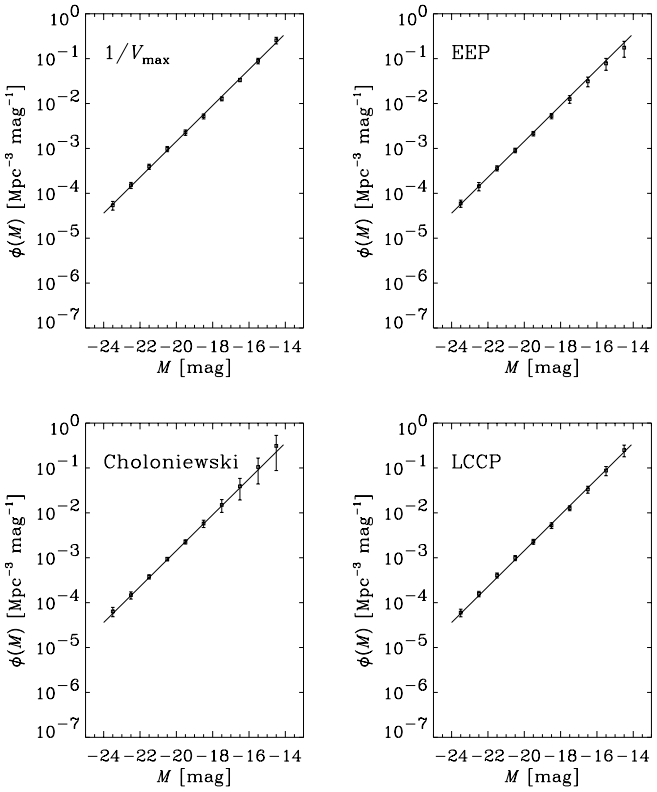

FIG. 6*a*

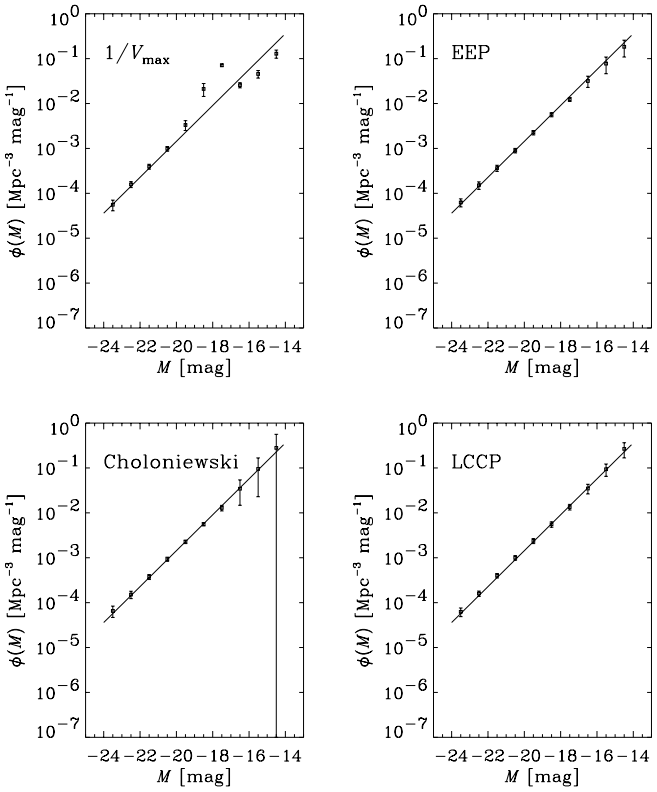

FIG. 6*b*

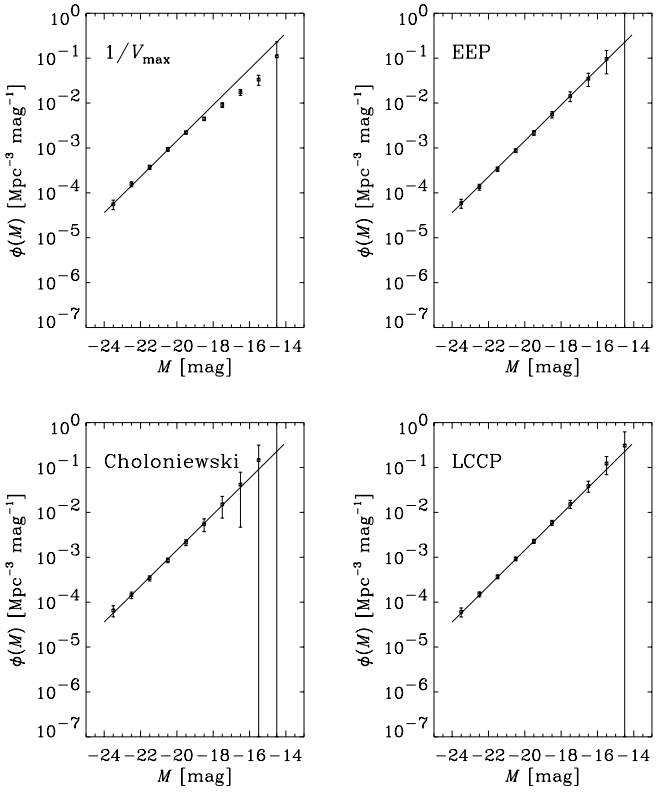

FIG. 6*c*

FIG. 6.—Same as Fig. 5, but for LF shape B

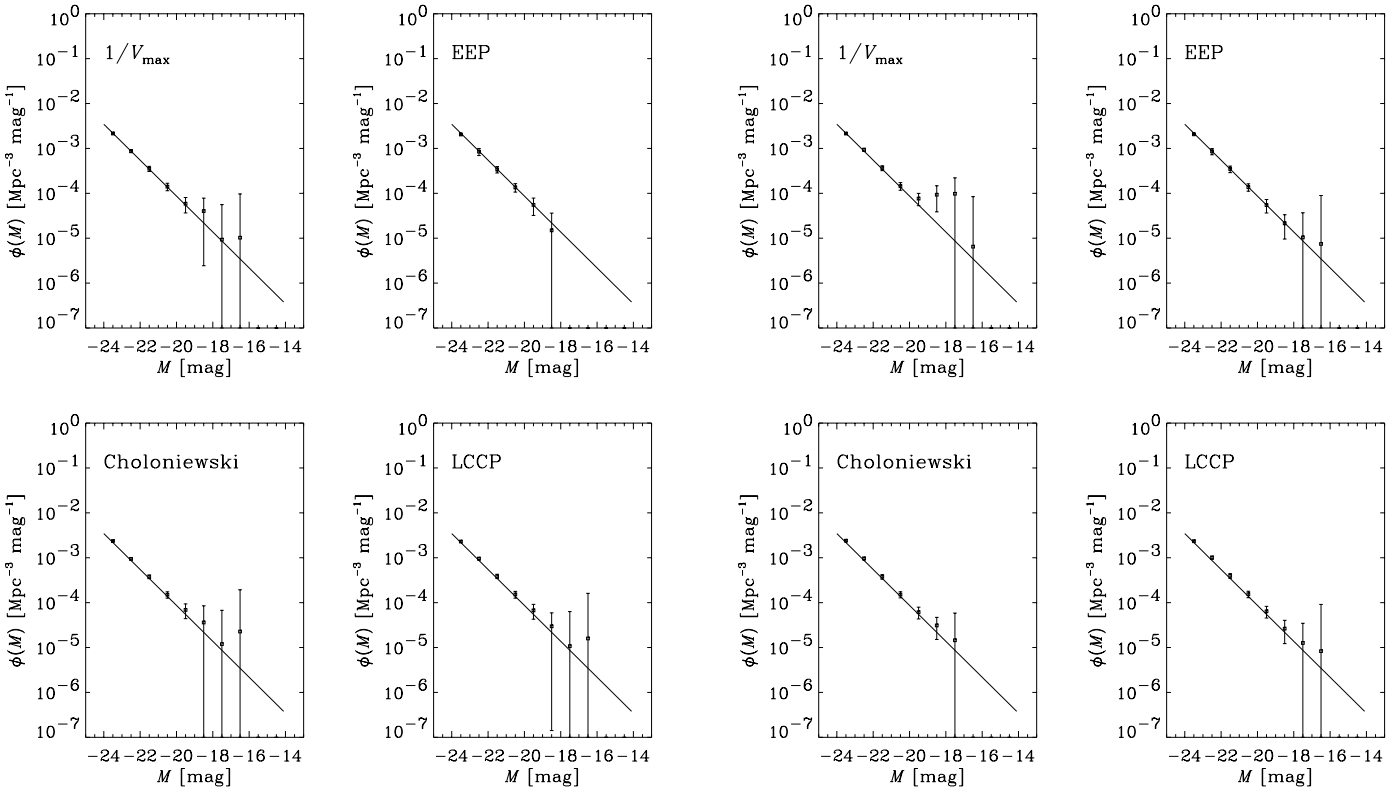

FIG. 7*a*

FIG. 7*b*

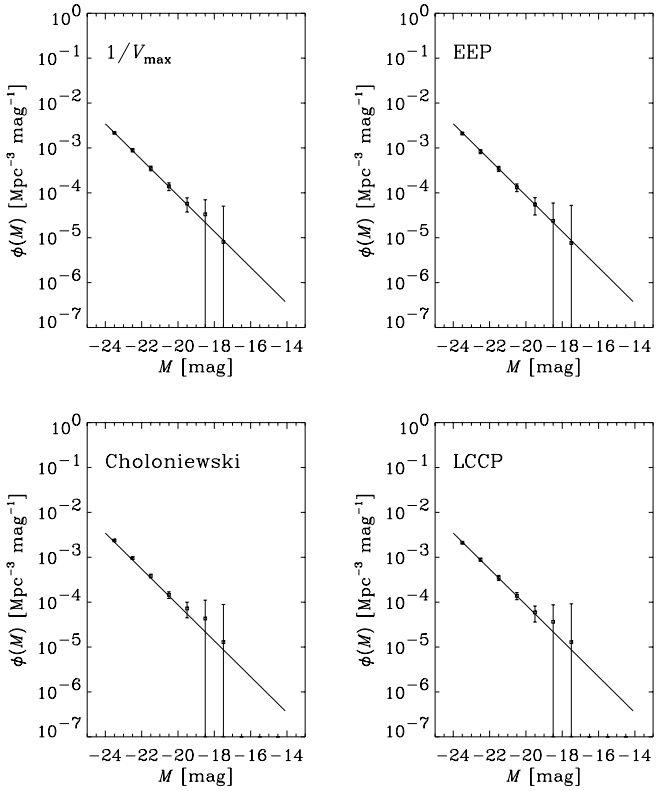

FIG. 7*c*

FIG. 7.—Same as Fig. 5, but for LF shape C

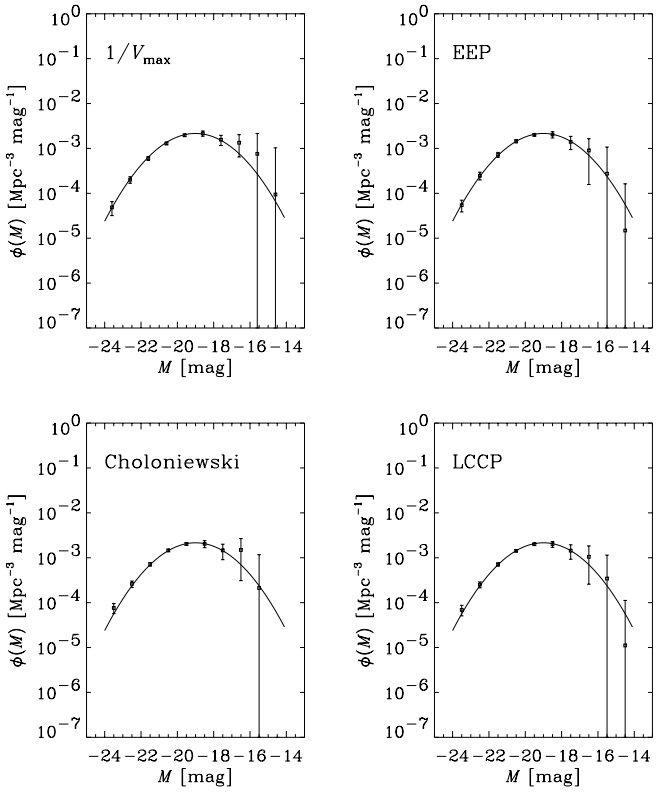

FIG. 8*a*

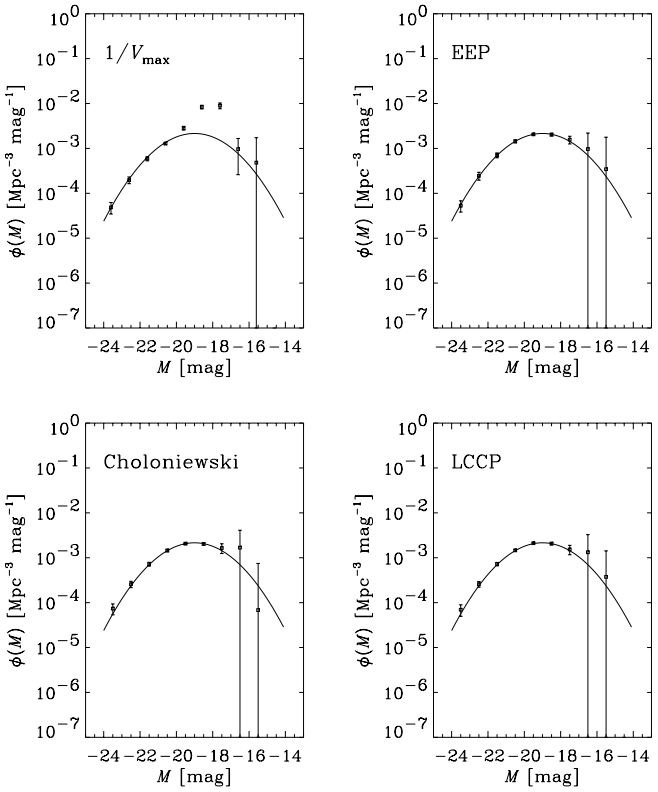

FIG. 8*b*

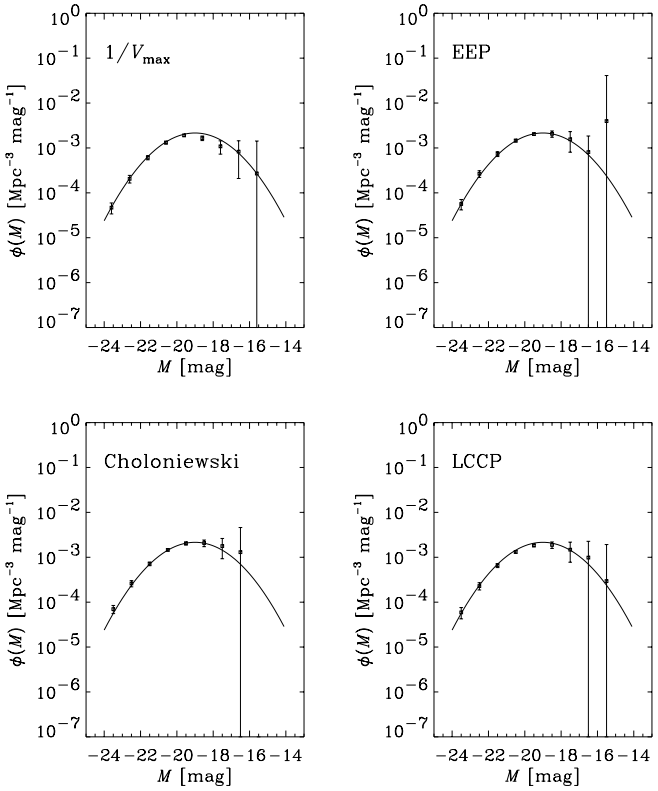

FIG. 8*c*

FIG. 8.—Same as Fig. 5, but for LF shape D

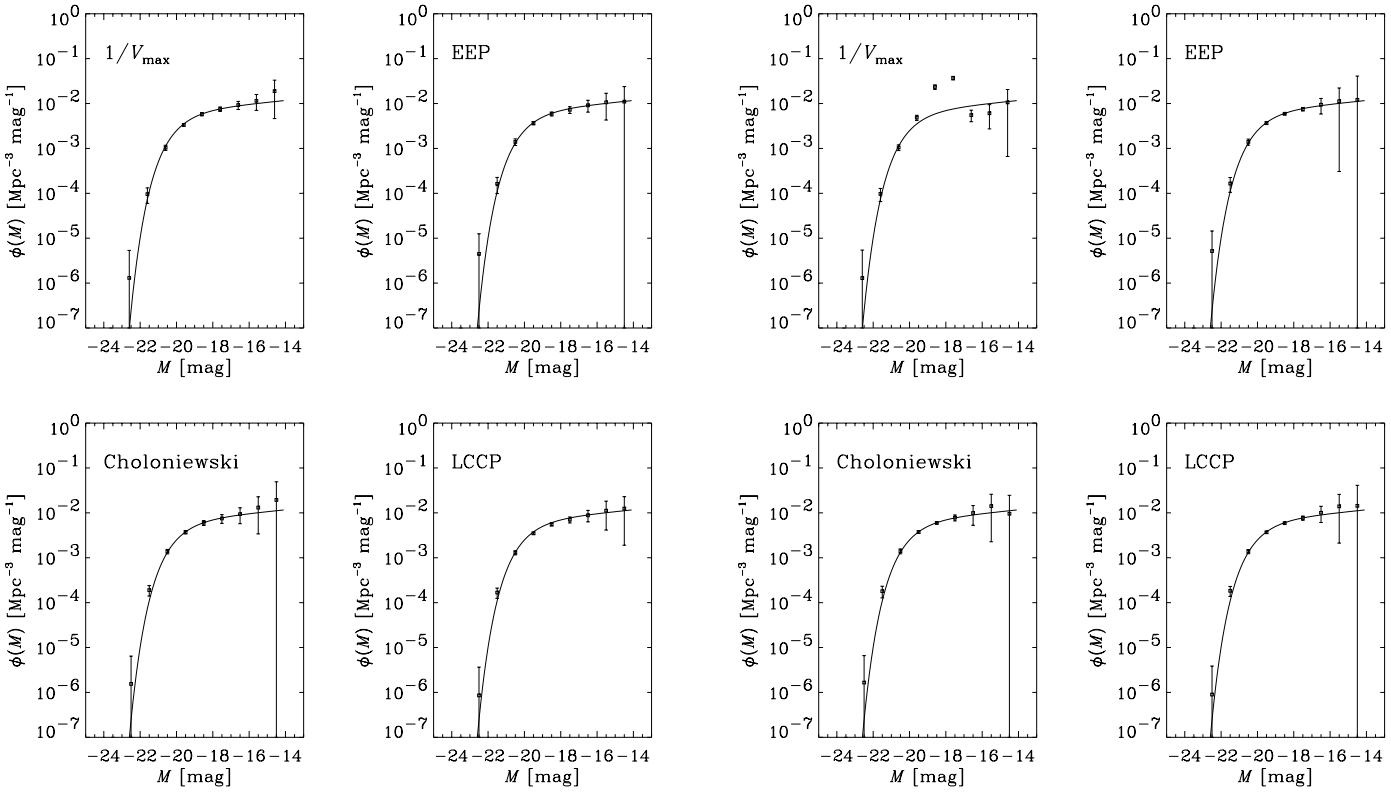

FIG. 9*a*

FIG. 9*b*

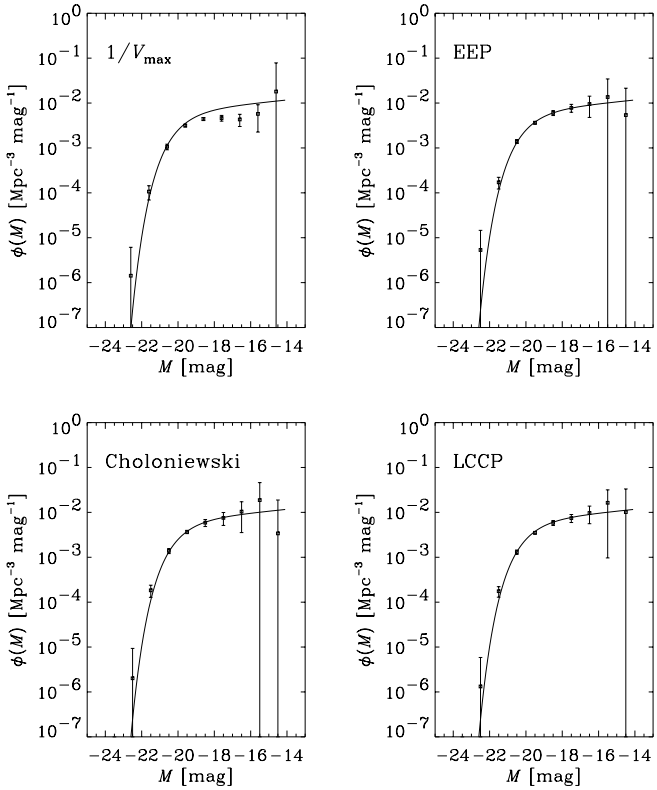

FIG. 9*c*

FIG. 9.—Same as Fig. 5, but for LF shape E

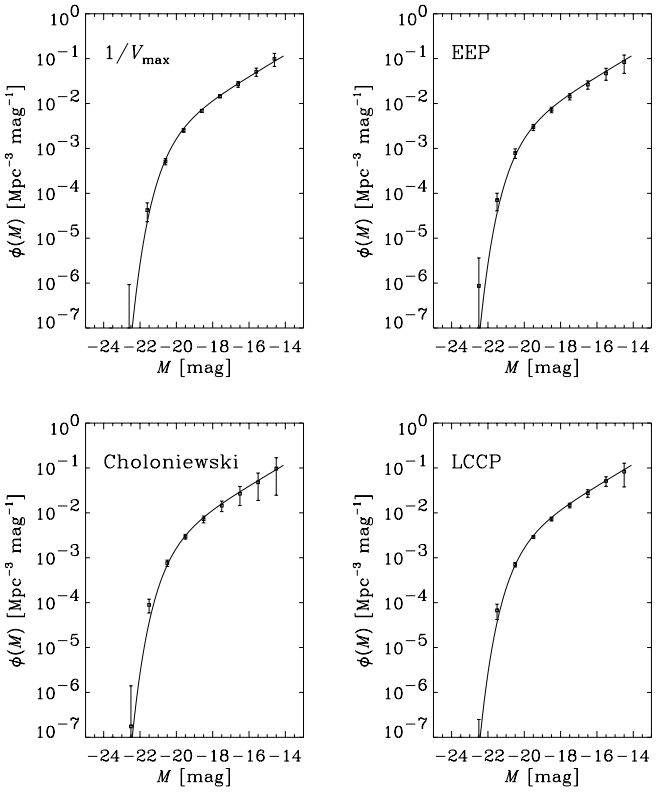

FIG. 10*a*

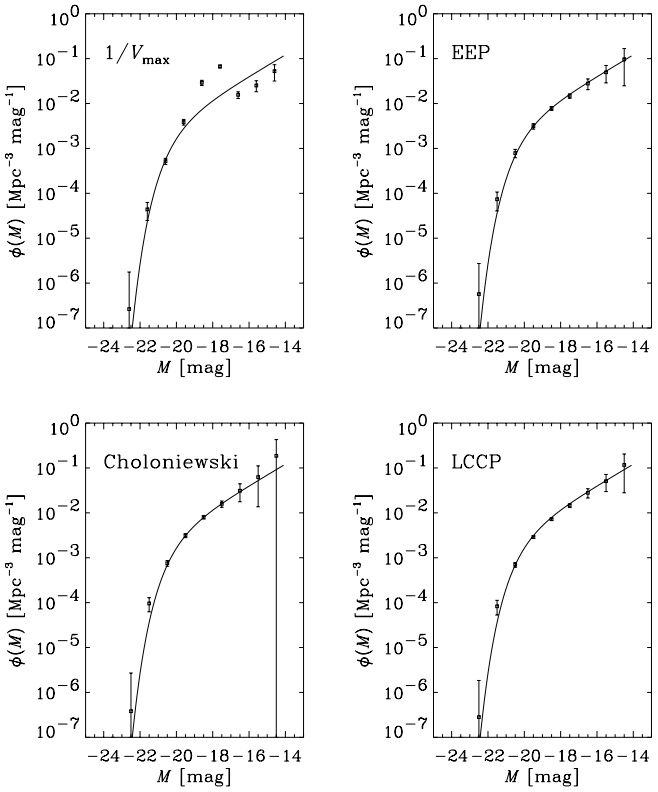

FIG. 10*b*

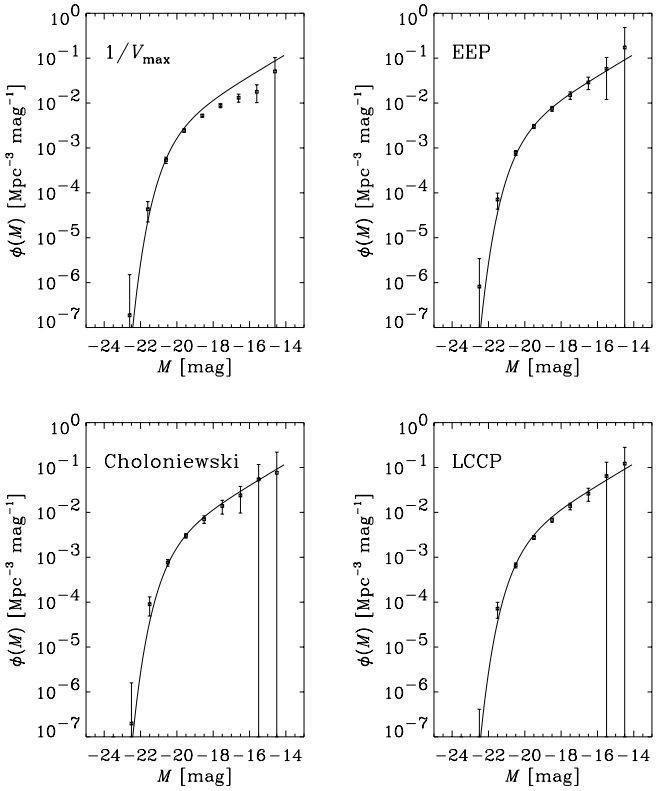

FIG. 10*c*

FIG. 10.—Same as Fig. 5, but for LF shape F

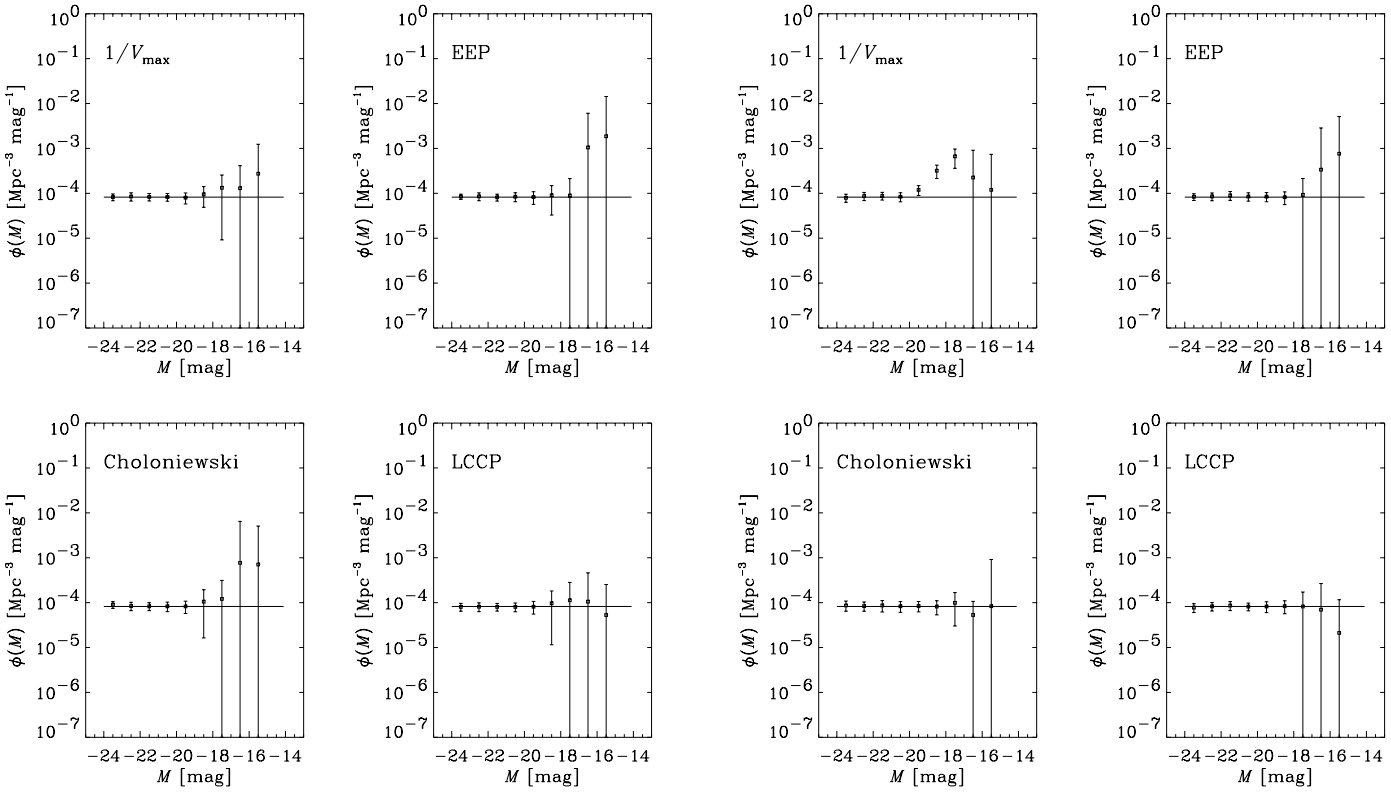

FIG. 11*a*

FIG. 11*b*

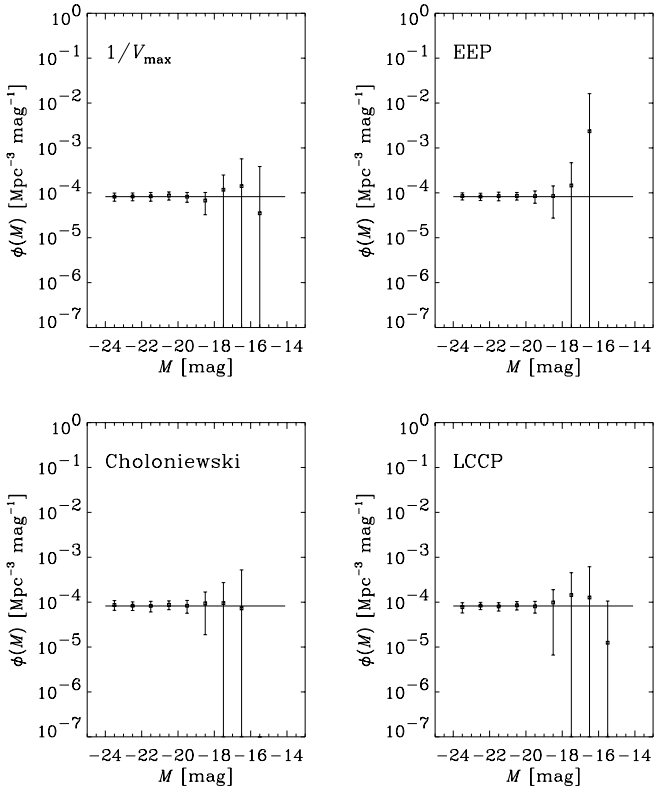

FIG. 11*c*

FIG. 11.—Estimates of the luminosity function (LF) for mock samples with size ~ 100. Shown for (*a*) the LF of the spatially homogeneous sample; (*b*) the LF from the cluster sample; and (*c*) the LF from the void sample. The input luminosity function is of the functional form A.

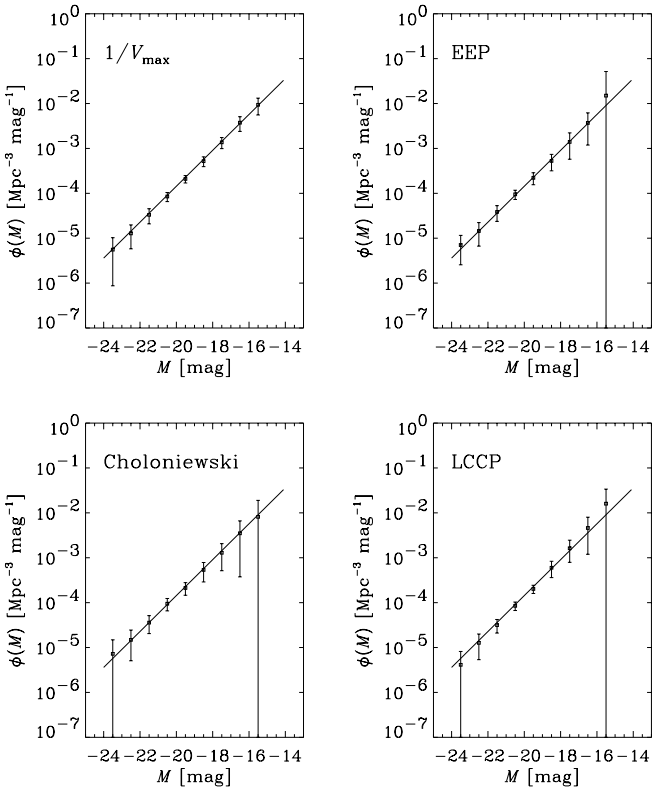

FIG. 12*a*

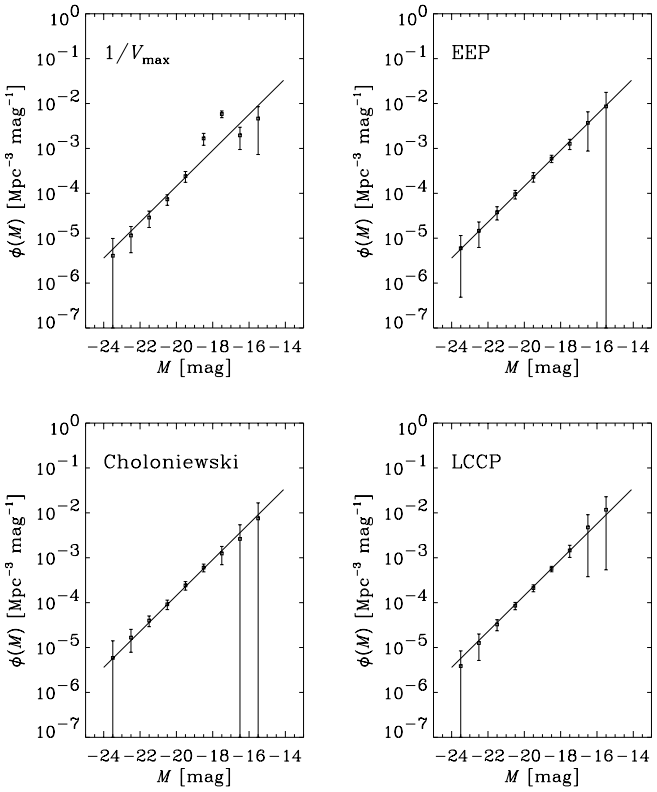

FIG. 12*b*

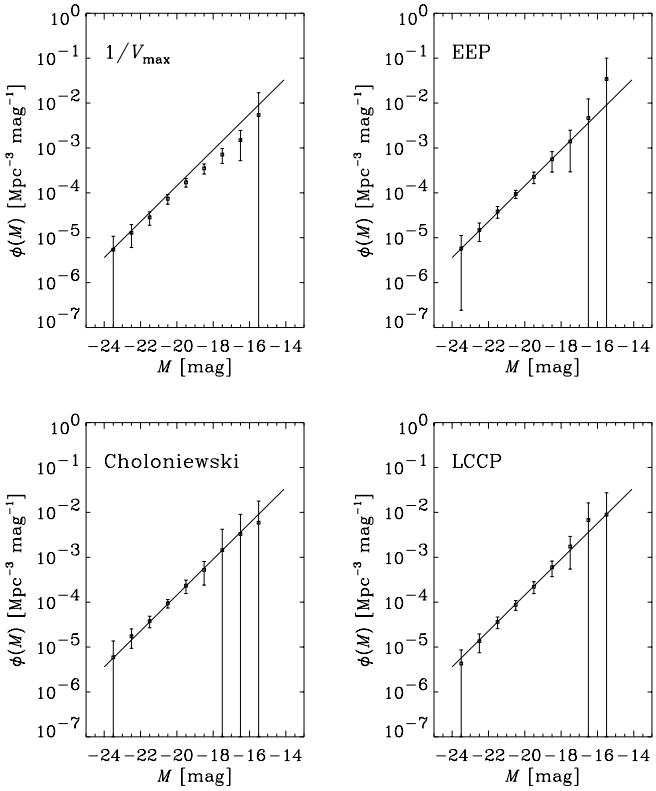

FIG. 12*c*

FIG. 12.—Same as Fig. 11, but for LF shape B

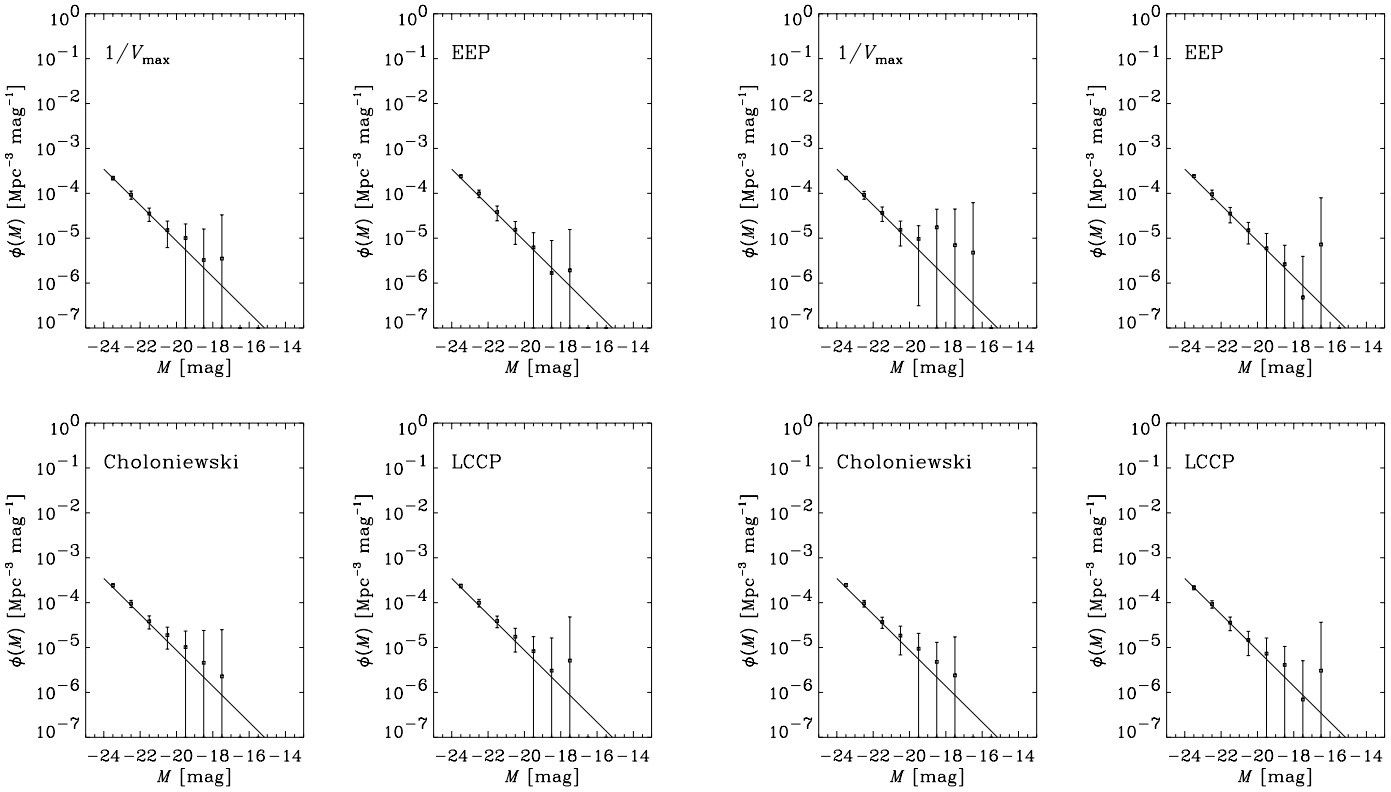

FIG. 13*a*

FIG. 13*b*

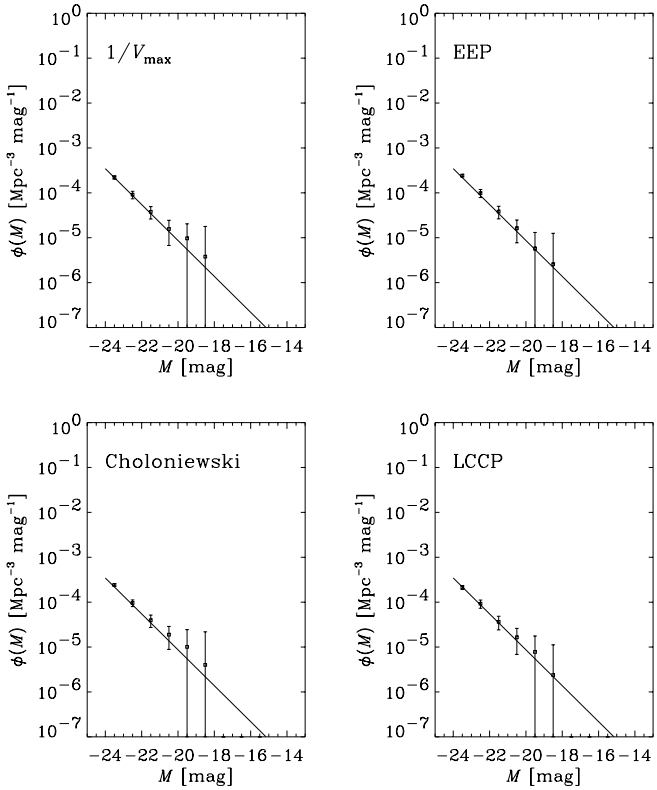

FIG. 13*c*

FIG. 13.—Same as Fig. 11, but for LF shape C

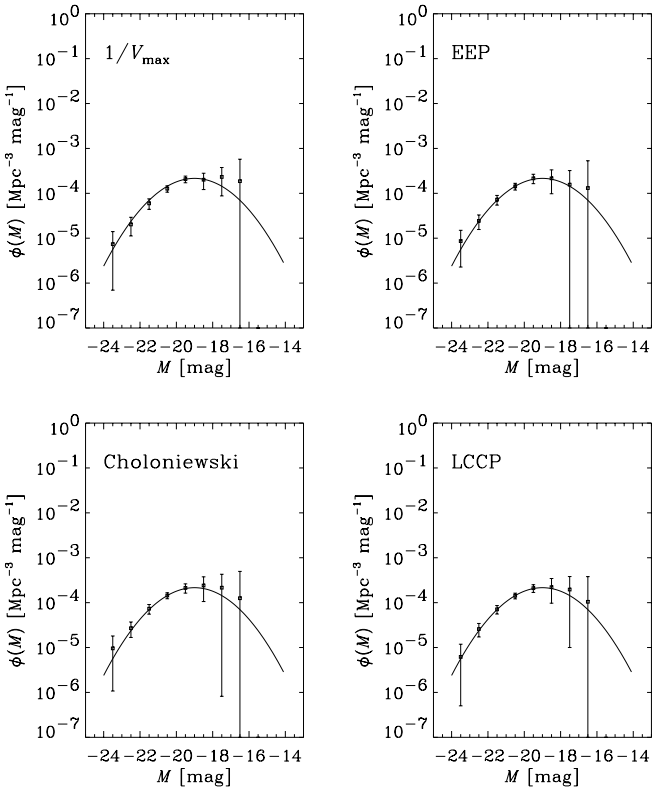

FIG. 14*a*

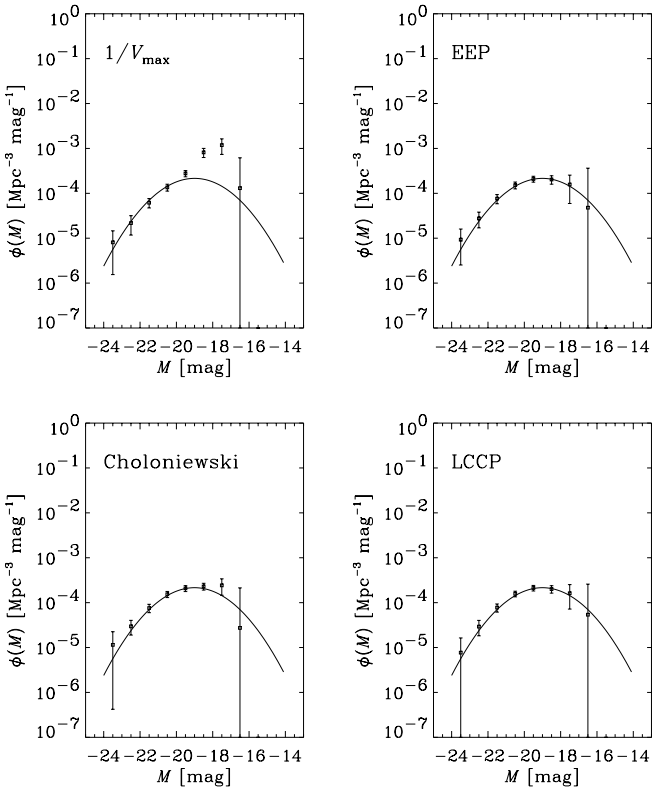

FIG. 14*b*

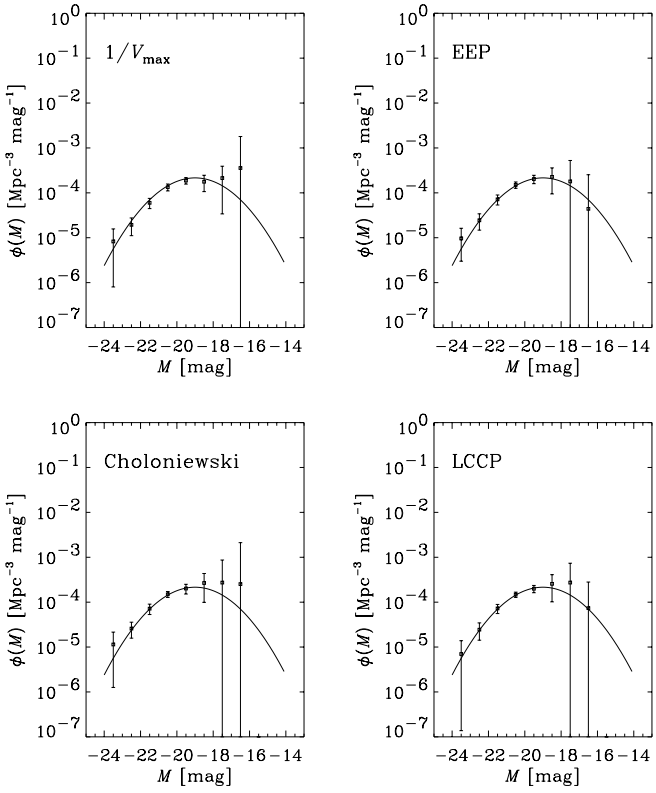

FIG. 14*c*

FIG. 14.—Same as Fig. 11, but for LF shape D

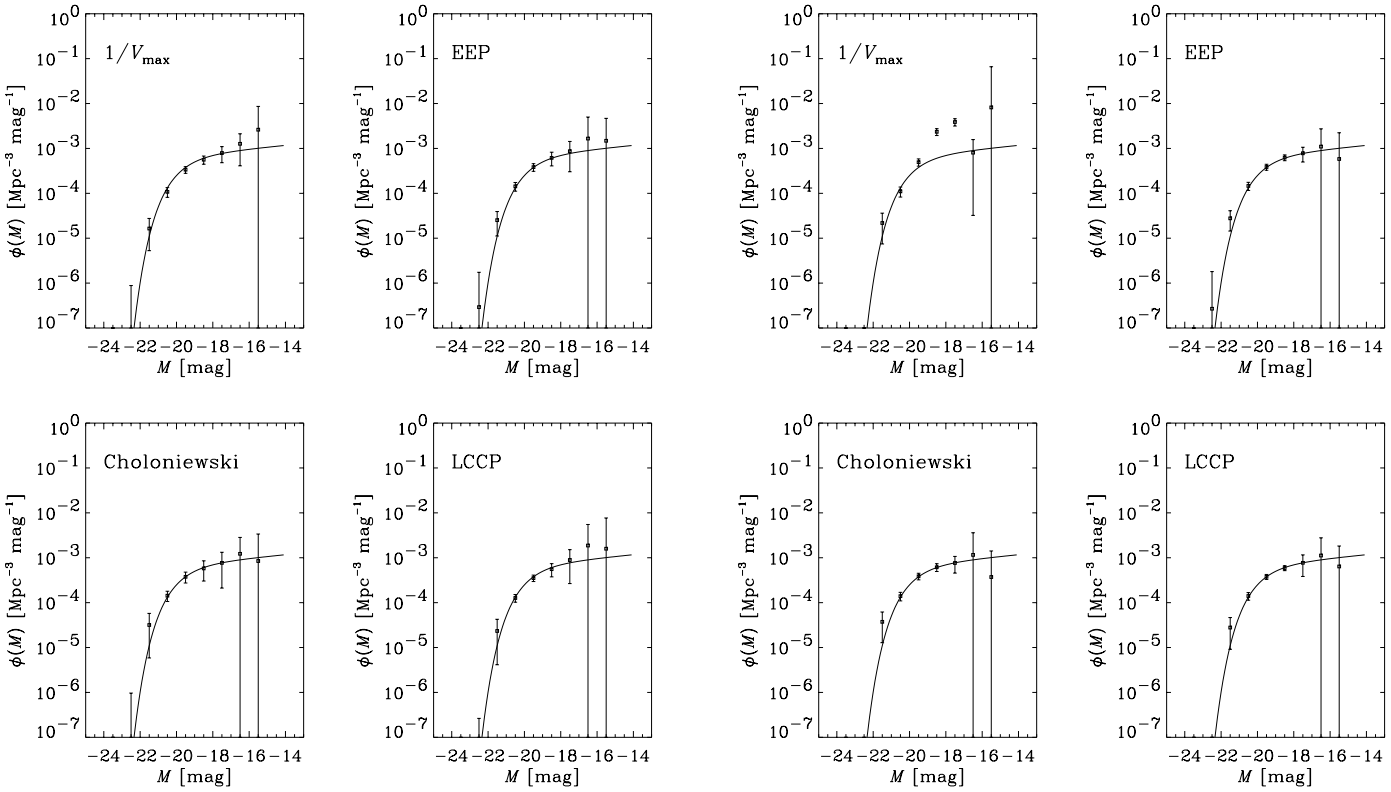

FIG. 15*a*

FIG. 15*b*

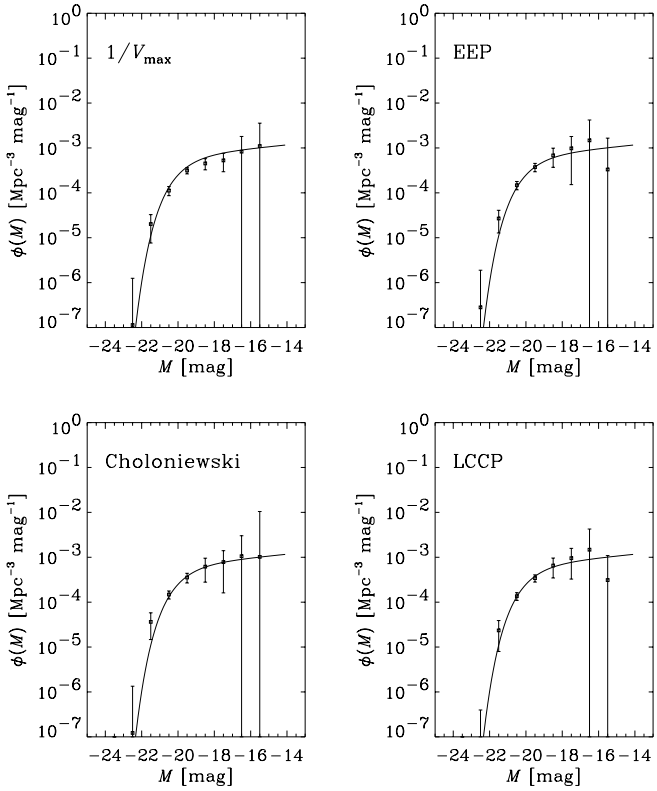

FIG. 15*c*

FIG. 15.—Same as Fig. 11, but for LF shape E

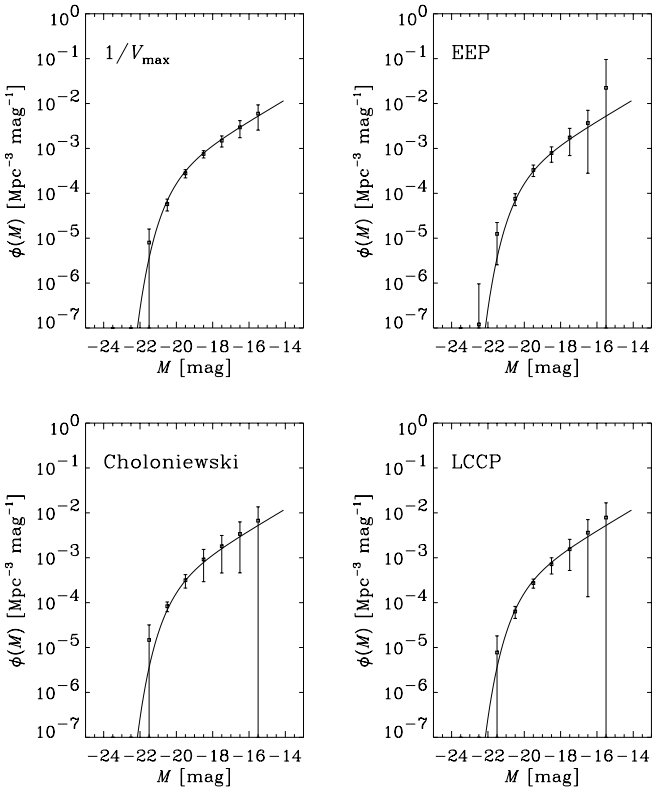

FIG. 16*a*

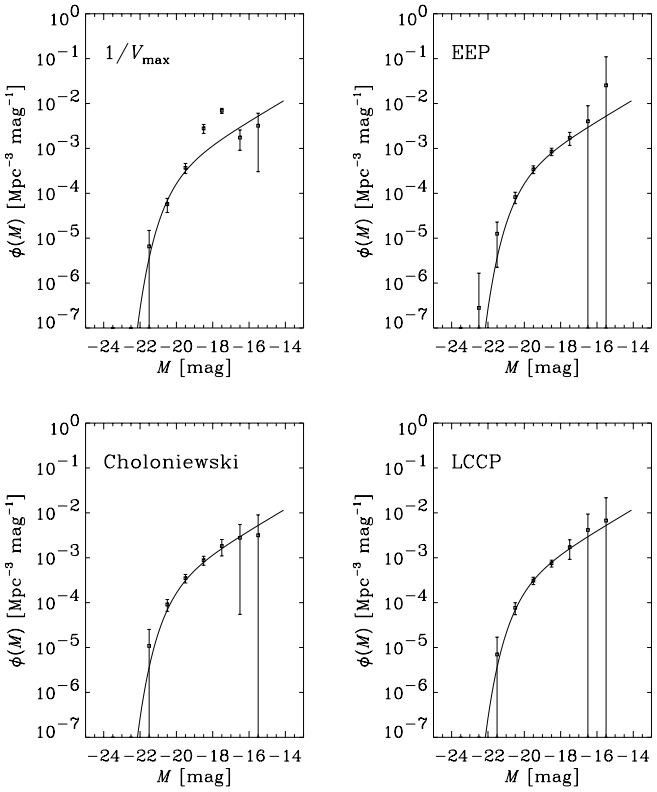

FIG. 16*b*

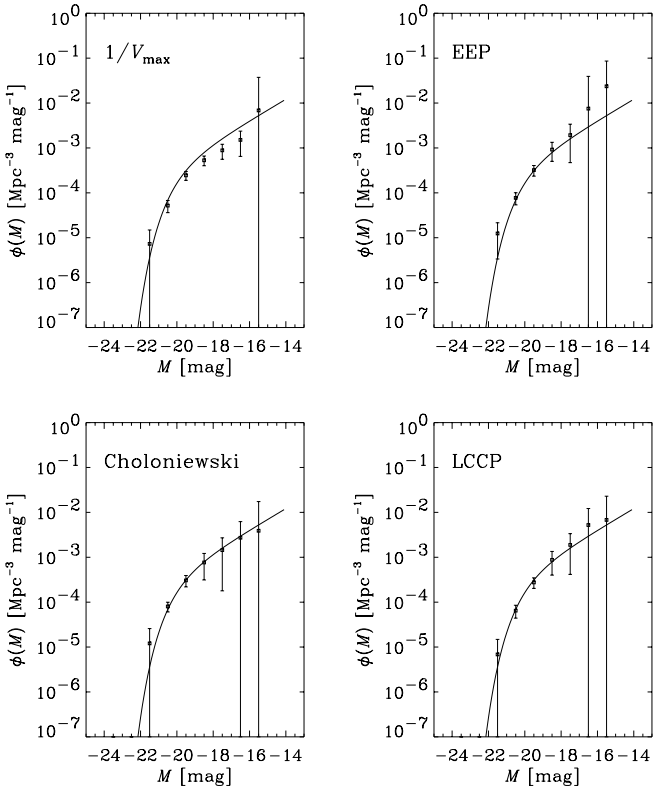

FIG. 16*c*

FIG. 16.—Same as Fig. 11, but for LF shape F

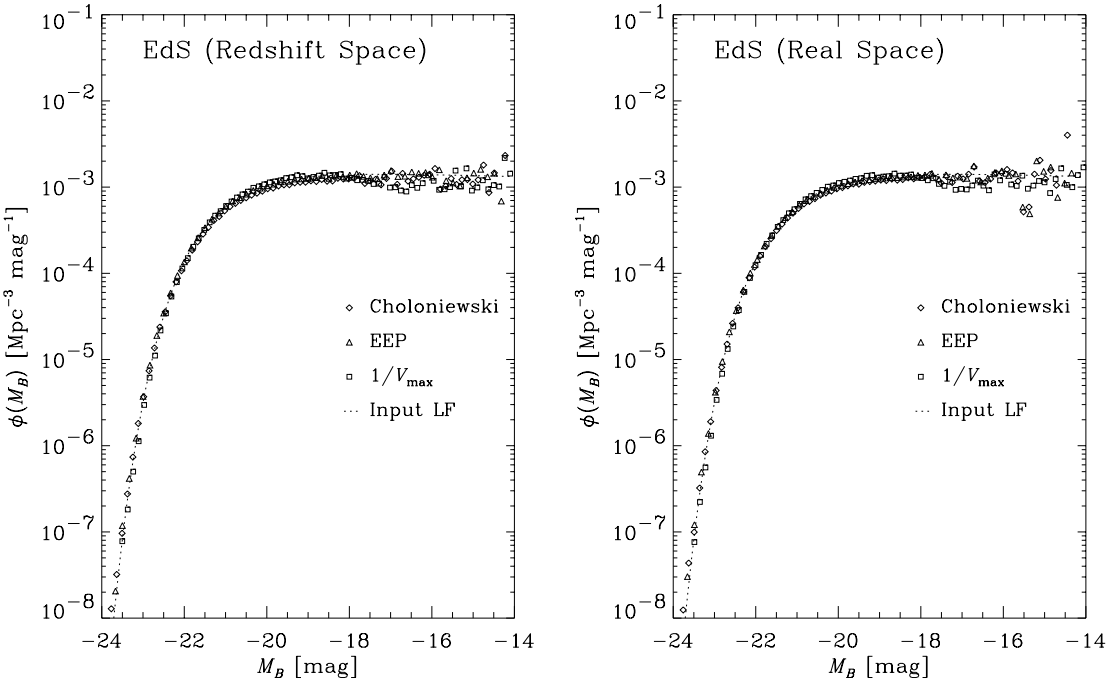

FIG. 17.—Input and estimated LFs of the mock 2dF catalog. The cosmological parameters are $\Omega_0 = 1.0$, $\lambda_0 = 0$, $\Gamma = 1$, and $\sigma_8 = 0.55$.

show the LFs derived from the redshift-space data, and the right panels show those from real-space data. First, we see that all the estimators provide perfectly consistent results, and they show an excellent agreement with the input LF. There are no significant difference between the real- and redshift-space data sets. The slight deviations of $1/V_{max}$ estimates are caused by clustering in the 2dF mock catalog. Thus, we do not need to consider the redshift distortion effect seriously when we derive the galaxy LF from such large-volume redshift surveys. When we use such a large survey, we should rather mention the photometric calibration as a more important error source.

4. APPLICATION TO THE HUBBLE DEEP FIELD

Recently, some authors have claimed that the faint-end slope of the LF becomes steeper with redshift at $z < 1$ (e.g., Ellis et al. 1996; Heyl et al. 1997; but see Lin et al. 1999). The LFs for some special classes of galaxies, such as

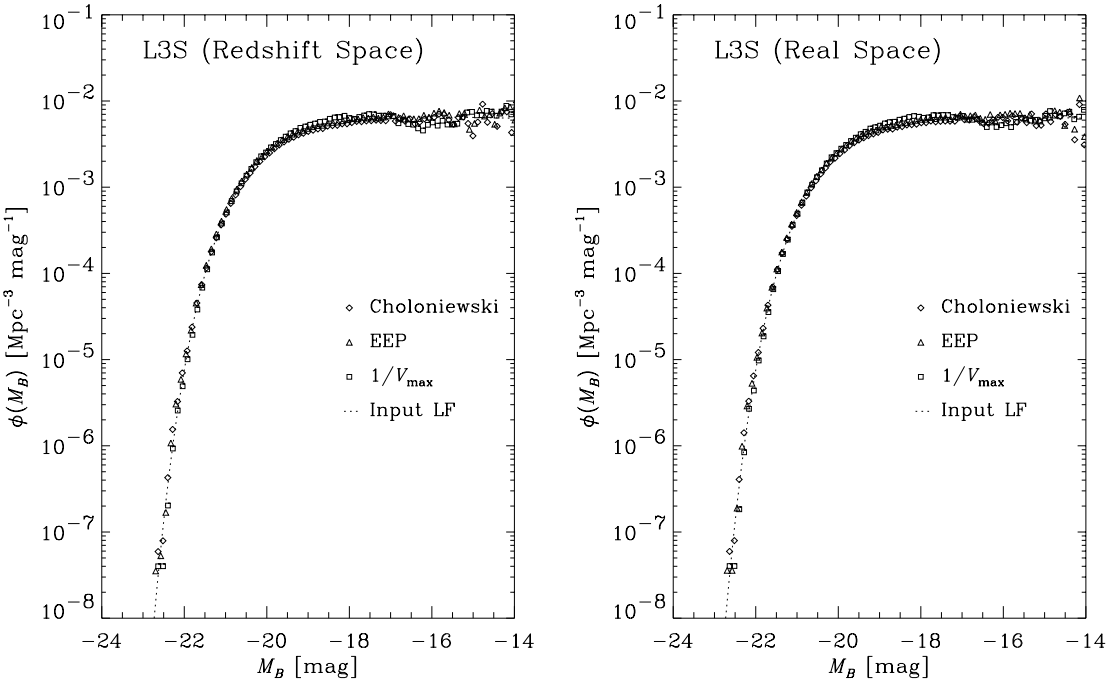

FIG. 18.—Same as Fig. 17, but for cosmological parameters $\Omega_0 = 0.3$, $\lambda_0 = 0.7$, $\Gamma = 0.25$, and $\sigma_8 = 1.13$

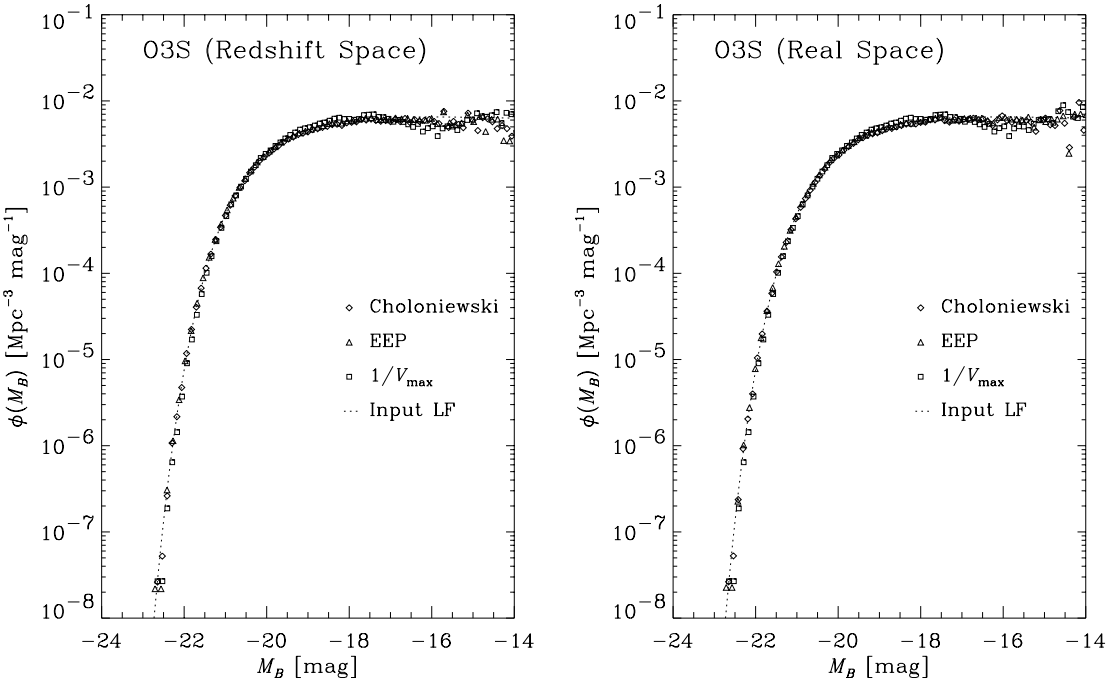

FIG. 19.—Same as Fig. 17, but for the cosmological parameters $\Omega_0 = 0.3$, $\lambda_0 = 0$, $\Gamma = 0.25$, and $\sigma_8 = 1.13$

Lyman-break objects (Steidel et al. 1998) or Lyα emitters (Pascarelle, Windhorst, & Keel 1998b), are now also available. However, we do not have a coherent understanding of the evolution of the LF and the evolution of the luminosity density, ρ_L. At low redshift, Zucca et al. (1997) reported a high-normalization LF with $\phi_* = 0.020\ h^3$ Mpc^{-3}, and Ellis et al. (1996) obtained $\phi_* = 0.026\,h^3$ Mpc^{-3}, while Loveday et al. (1992) derived $\phi_* = 0.014\ h^3$ Mpc^{-3}, and the Las Campanas Redshift Survey result (Lin et al. 1996) is similar to the value of Loveday et al. (1992). The local value of the LF parameters plays a crucial role in the study of galaxy evolution, since it controls the redshift dependence of ρ_L. Cowie, Songaila, & Barger (1999) showed a rather mild evolution of the UV luminosity density at $z < 1$ from their surveys. On the other hand, high-redshift LF estimations are also controversial with each other. Gwyn & Hartwick (1996) claimed dramatic changes in the LF from $z = 0$ to $z \sim 5$, becoming flat between $-24 \leq M_B \leq -15$ for $3 < z < 5$. On the other hand, SLY97 reported a more familiar Schechter form with $\alpha = -1.3$ for the LF at $3 < z < 4$. Mobasher et al. (1996) suggested a stronger evolution of the LF. From a deep multiband photometric survey, Bershady et al. (1997) gave a constraint that ruled out Gwyn & Hartwick's result.

In this section, we apply the four estimators to the photometric redshift catalog of the HDF to study the evolution of the LF shape. For the observational data, the error estimation is complicated, because the estimation procedure for the LF involves magnitude selection, weighting, etc. In such cases, bootstrap resampling analysis is known to be often superior to classical analytic methods in estimating statistical properties (e.g., Efron & Tibshirani 1993; Babu & Feigelson 1996; Davison & Hinkley 1997). Thus, we used the bootstrap method for estimating the statistical uncertainties. When we perform the bootstrapping, how we generate good random numbers is important. We generated the uniform random number by the Mersenne twister method,[7] developed by Matsumoto & Nishimura (1998).

4.1. *Sample*

We use the photometry and photometric redshift catalog of the HDF prepared by FLY99. Their catalog contains 1067 galaxies, with AB(8140) < 26.0. The photometric redshifts are derived based on both *UBVI* (F300W, F450W, F606W, and F814W; Williams et al. 1996) obtained by WFPC2, and *JHK* obtained by the IRIM camera on the Kitt Peak National Observatory (KPNO) 4 m telescope. The object detection and photometry are performed using SExtractor (Bertin & Arnouts 1996). Details of the procedures are found in FLY99. In the peripheral region of the WFPC2 image (referred to as zone 2), the detection limit is AB(8140) = 26 mag, and in the inner region (zone 1), AB(8140) = 28 mag. We restrict our analysis to the inner zone 1 sample. The solid angle of zone 1 is 3.92 arcmin^2 = 3.32×10^{-7} sr. The sample size is then 946 galaxies.

FLY99 used four spectral templates given by Coleman, Wu, & Weedman (1980) to determine the photometric redshifts. The templates are extrapolated for ultraviolet wavelengths using the results of Kinney et al. (1993), and for infrared using the models of Bruzual & Charlot (1993). Evolutionary corrections are not included in the model spectra to avoid additional parameter dependence. According to Coleman et al. (1980), they classified the galaxy spectra into four categories: (1) Elliptical, (2) Sbc, (3) Scd, and (4) Irr. We used these labels to set the *K*-corrections.

In principle, the SED must be the same as the templates used in FLY99, but for simplicity and comparison with other studies, we used the galaxy SED sample compiled by

[7] For recent development, see http://www.math.keio.ac.jp/matumoto/mt.html.

Kinney et al. (1996). The data of Kinney et al. (1996) have almost the same properties as the SED templates of FLY99; thus, we can use them comfortably. To construct the K-correction function, we first selected the sample galaxy SEDs corresponding to the labels of FLY99, and then fitted polynomial functions from 1st order to 6th order. The order of polynomial fitting was decided by referring to AIC, and we chose the 5th order.

4.2. *Results and Discussions*

We show the redshift-dependent LF at the I band and the B band in Figures 20 and 21, respectively. The symbols represent the LFs estimated by the four methods. We show the LFs of the HDF at $0 < z < 0.5$ (106 galaxies), $0.5 < z < 1.0$ (193), $1.0 < z < 1.5$ (204), $1.5 < z < 2.0$ (193), $2.0 < z < 3.0$ (117), and $3.0 < z < 6.0$ (109). The sample is

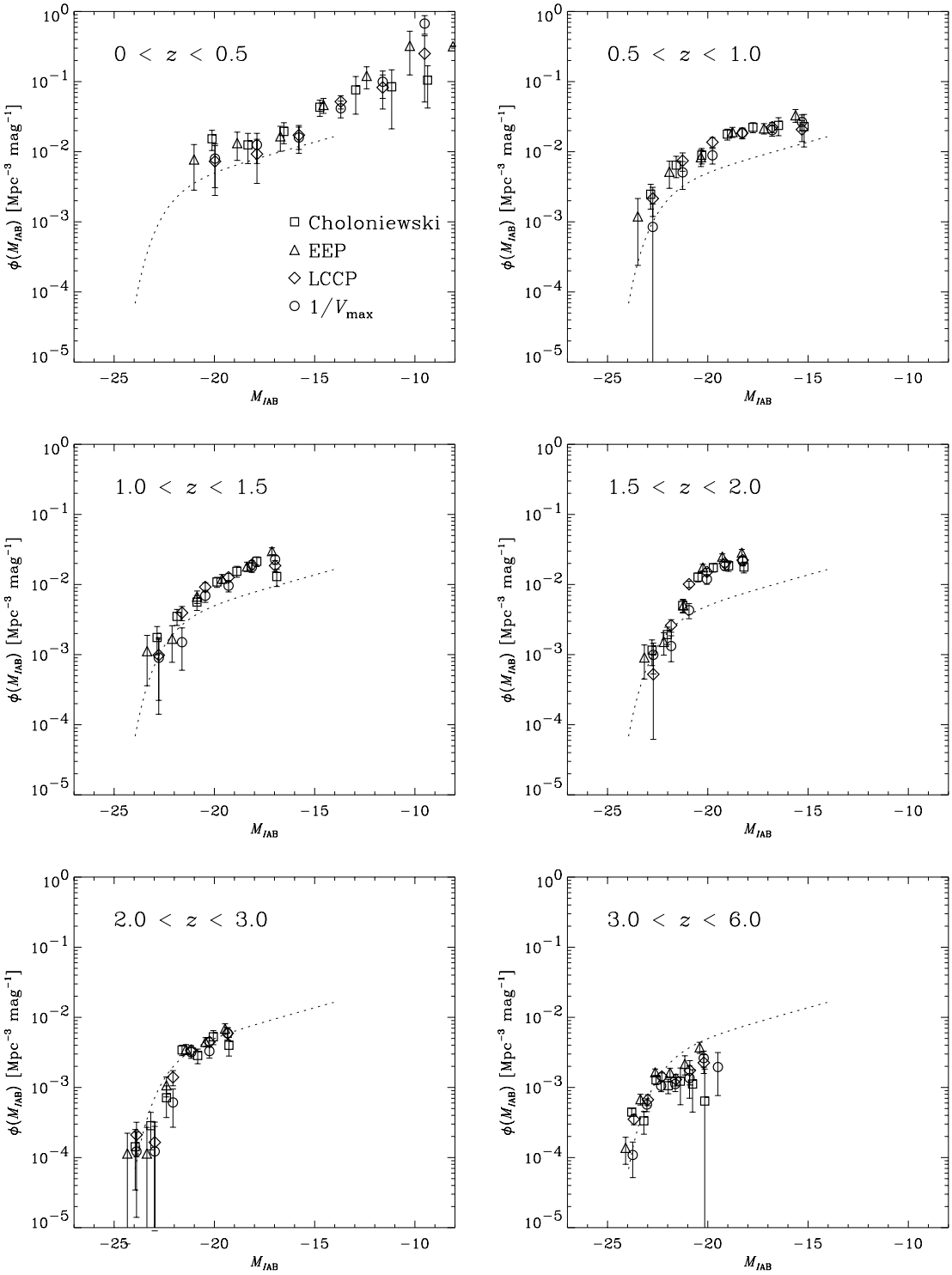

FIG. 20.—I-band LF derived from the HDF photometric redshift catalog prepared by Fernández-Soto et al. (1999). The dotted line represents the local I-band LF obtained by Metcalfe et al. (1998).

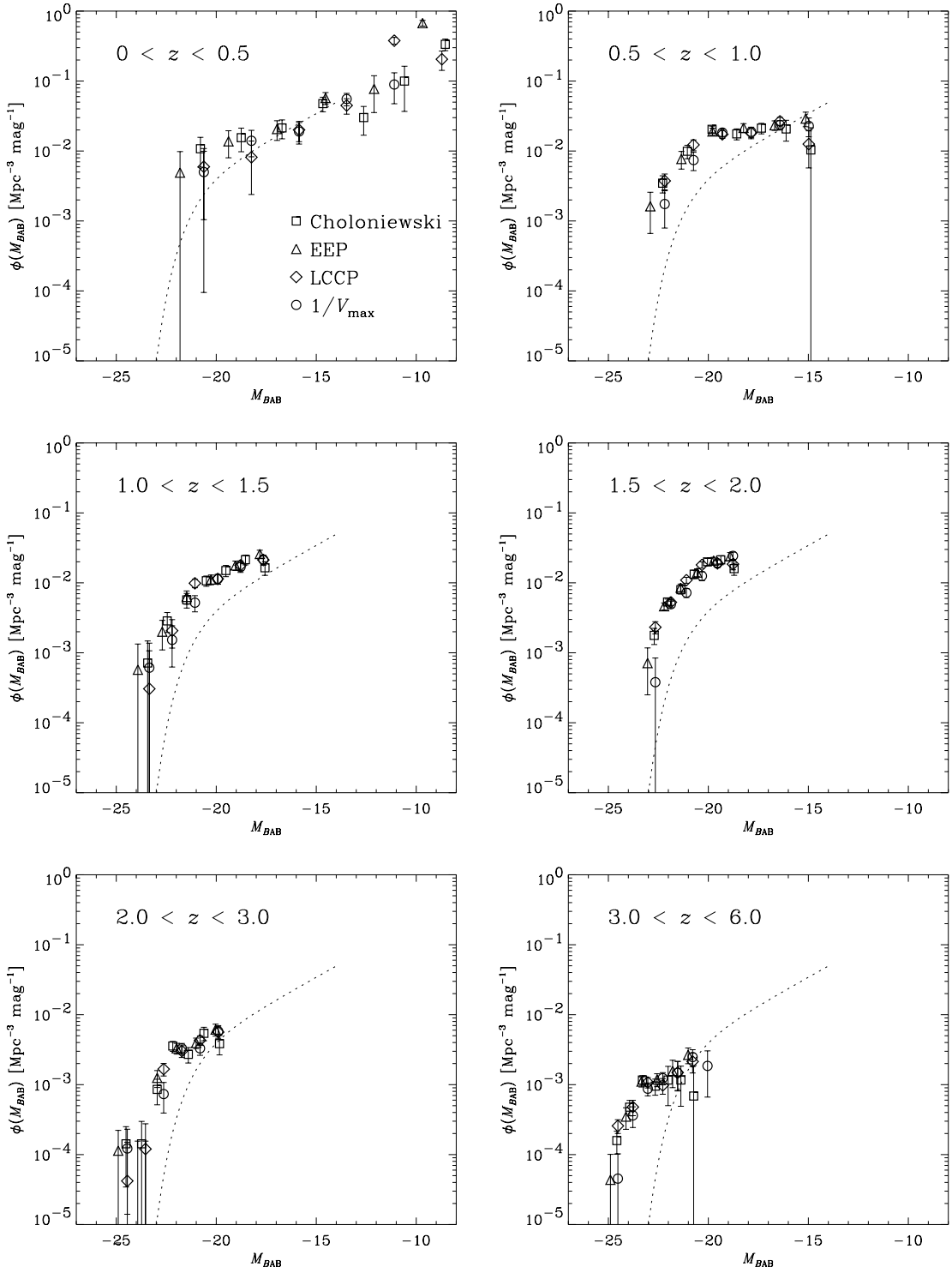

FIG. 21.—Same as Fig. 20, but for the *B*-band LF. The dotted lines show the *B*-band LFs derived by Sawicki et al. (1997).

I-band–selected, and we derived the *B*-band LF by following the discussion of Lilly et al. (1995). We stress that the four different LF estimators give consistent results for the HDF sample, being the same as the results for the mock catalogs.

We clearly see the evolutionary trend of the LF with redshift. However, we note that although we can fit the Schechter function, it is not so easy to derive the parameters α, M_*, or ϕ_* precisely, because the Schechter function is rather smooth and the errors of these characteristic parameters are strongly correlated. These parameters can be easily affected by statistical fluctuations. We discuss more details of the *I*- and *B*-band results at each redshift range in the following subsections.

4.2.1. *I-Band LF at* $0 < z < 0.5$

In Figure 20, the dotted line represents the local *I*-band LF obtained by Metcalfe et al. (1998). Metcalfe et al. (1998)

pointed out a possible upturn of the faint end of their multiband LFs, although they took a prudent attitude in concluding firmly. The upturn magnitude, $M_I \sim -15 + \log h$ mag (in Fig. 20, $h = 0.75$), is in good agreement with that of our lowest redshift LF, except for the normalization.

4.2.2. *Evolution at the B-Band:* $0 < z < 0.5$

We compare the normalization of the LF with other previous results. Our B-band LF shows roughly good agreement with other local LFs. In Figure 22, we plot our LF, the SLY97 Schechter fit, and Schechter functions reported by Metcalfe et al. (1998) and Ellis et al. (1996; Autofib Redshift Survey). The dotted line depicts the SLY97 LF, the dot-dashed line represents the Metcalfe et al. (1998) B-band LF, and the long-dashed line shows the Autofib LF at $z < 0.1$. Our LF and that of SLY97 agree with the higher normalization LF reported by the Autofib Survey, but they are significantly higher than that of Metcalfe et al. (1998). The Autofib LF is also consistent with the LF of the ESP Survey (Zucca et al. 1997), while Metcalfe et al.'s LF is consistent with the EEP LF and the Stromlo-APM LF (Loveday et al. 1992). However, since the solid angle covered by the HDF is extremely small and thus the normalization can be strongly affected by cosmic variance, we should not go into any further discussion.

We note that the SLY97 M_* value is significantly higher than those of other surveys. This is because the exponential decline at the bright end is not observed in the HDF LF at $0 < z < 0.5$, and a bump exists at $M_B \sim -20$ mag. We should also mention that the error bar of the M_* of SLY97 is very large (1.6 mag). Considering the large error bars and the uncertainty of the photometric redshift, we conclude that the bright M_* is not a real feature.

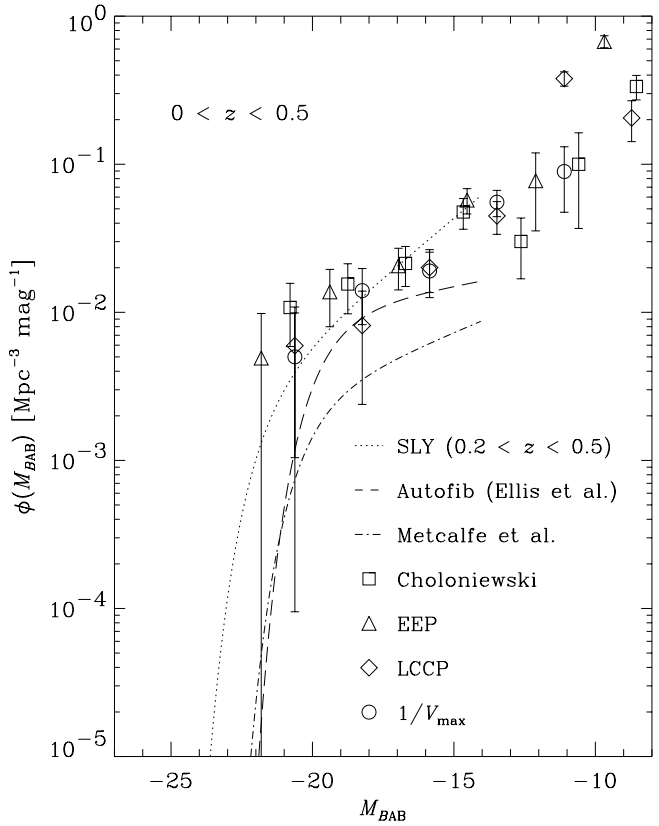

FIG. 22.—B-band LF of the HDF in the redshift range $0 < z < 0.5$. Dotted line depicts the Sawicki et al. (1997) LF, dot-dashed line represents the Metcalfe et al. (1998) B-band LF, and long-dashed line shows the Autofib LF at $z < 0.1$. Our LF and that of Sawicki et al. (1997) agree with the higher normalization LF reported by the Autofib Survey, but are significantly higher than that of Metcalfe et al. (1998).

At this lowest redshift, the rise of the faint end is prominent. The problem of the faint-end slope of the galaxy LF has long been a matter of debate, and we do not yet have a widely accepted consensus. As we already pointed out above, even in the I-band we find a steepening of the faint end. If this steep faint end is an artifact of clustering, then the LF derived from $1/V_{max}$ and those derived from other estimators should have been different (§ 3.1). In fact, they are consistent with each other. Thus we conclude that, at least in the HDF, the faintest end of the LF has a steep slope in the local universe.

4.2.3. *Evolution at the B-Band:* $0.5 < z < 1.0$

It seems that the brighter galaxies are more numerous than the local value at $0.5 < z < 1.0$. Here we should remember that the "fuzzy" redshift determination is known to affect the shape estimation (Liu et al. 1998). Liu et al. (1998) showed by numerical experiments that the faint-end slope is underestimated and M_* is overestimated by the photometric redshift blurring. The uncertainty of the photometric redshift is rather independent of the object redshift, so the effect will be severer at low z, and M_* can be overestimated. Thus, the increase of bright galaxies is partially due to this effect. However, we can discuss the trend of the LF evolution by comparing the LF derived from photometric redshifts consistently (Liu et al. 1998).

4.2.4. *Evolution at the B-Band:* $1.0 < z < 2.0$

The LFs of the redshift ranges $1.0 < z < 1.5$ and $1.5 < z < 2.0$ are the most reliable ones among the LFs shown in Figure 21, since the sample size is twice as large as those of the other redshift ranges, and in addition, the photometric redshift error becomes worse again at $z > 2$. SLY97 suggested a steepening of the faint-end slope at this redshift. Our LF of $1.5 < z < 2.0$ presents a similar feature, although the slope becomes flatter at the faintest regime. The deformation of the LF from $z \sim 0$ to $z \sim 2$ supports the idea that the steepening of the fainter side of the LF, which is confirmed at $z < 1$, is continued up to $z \sim 2$. We do not find a significant shift of M_* at this redshift range.

4.2.5. *Evolution at the B-Band:* $2.0 < z$

The normalization of the farthest redshift LFs settles down to the local value, while we also find a brightening of M_* at $z > 3.0$. We must be careful because at such high redshifts, cosmological surface brightness dimming is quite severe, and the selection effect becomes significant (Ferguson 1998; Weedman et al. 1998). Other kinds of selection effects are discussed in Pascarelle et al. (1998a). Therefore, there can exist more numerous galaxies than estimated. Further discussions require delicate treatment of such effects.

4.2.6. *Luminosity Density Evolution*

In order to explore the cosmic star formation history, we derive the luminosity density at the B- and I-bands based on our LFs. We fit the Schechter function and extrapolate the faint end below the detection limit. As we mentioned above, the Schechter parameters are poor indicators of the galaxy evolution, but the integrated luminosity density, ρ_L,

is regarded as an indicator of the evolution of galaxies, because while the Schechter parameters are significantly affected by the fluctuations, ρ_L is robust against the effect. We show the derived ρ_L in Figure 23. The top panel shows the evolution of the B-band luminosity density, $\rho_L(B)$, and the bottom shows the I-band luminosity density, $\rho_L(I)$. Open squares show $\rho_L(B)$ derived from CFRS (Lilly et al. 1995); open circles, $\rho_L(B)$ from Autofib (Ellis et al. 1996); the open triangle represents the value from the Stromlo-APM (Loveday et al. 1992); and the open diamond, the ESP value (Zucca et al. 1997). Crosses show the estimates of SLY97. In this paper we did not try to correct for the reddening effect of dust.

We see the local diversity of the $\rho_L(B)$, corresponding to the normalization discrepancy in Figure 23. Despite the fact that the local LF is difficult to derive from the HDF data, our low-z value is consistent with other previous results. In addition, our $\rho_L(B)$ at $0.5 < z < 1.0$ suffers significantly from the redshift-blurring effect, but it is also consistent with the CFRS highest redshift point within the errors. As a whole, our result is consistent with that of SLY97, except for $2.0 < z < 3.0$. In this redshift range, the $\rho_L(B)$ of SLY97 is several times larger than our estimate. This difference may be because SLY97 obtained a steeper α and brighter M_* than ours. We find a flatter LF slope, and the estimates fainter than -20 mag are not reliable in our result, since the fluctuation is horribly large at this redshift. If we choose a steeper slope, our ρ_L will be higher. We must wait for larger data sets to address this problem. At very high z, we derive a moderately high ρ_L, implying that significant numbers of stars have already formed at such a high redshift. The evolution of $\rho_L(I)$ appears to be flat. At the longer wavelength, the observed light is dominated by the contribution from lower mass stars, and the temporal change of the star formation rate (SFR) is less prominent. We should also note that the I-band results are subject to larger K-correction extrapolation uncertainties compared to the B-band results.

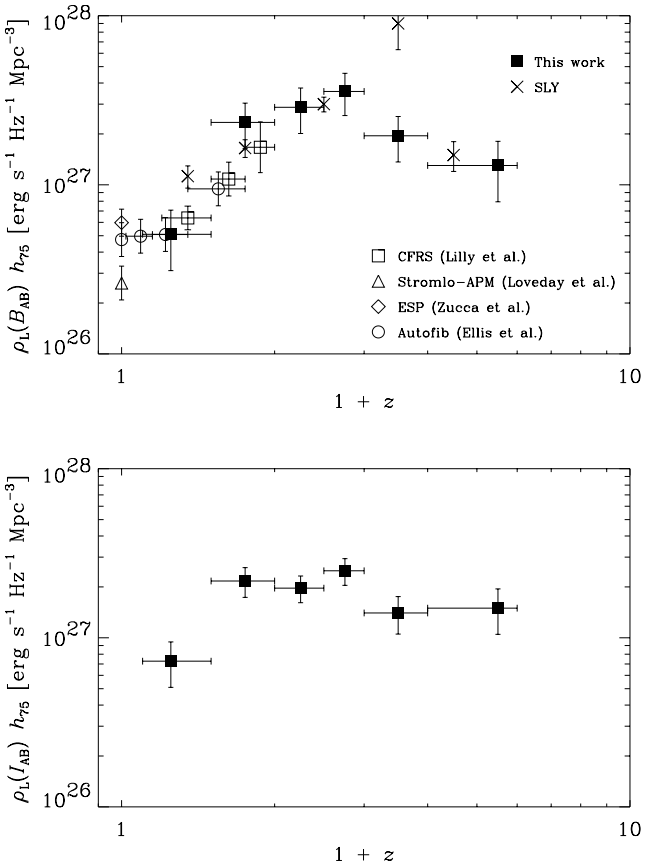

FIG. 23.—Evolution of the luminosity density, ρ_L. *Top*: B-band luminosity density. *Bottom*: I-band luminosity density. Open squares show $\rho_L(B)$ derived from CFRS (Lilly et al. 1995), open circles show $\rho_L(B)$ from Autofib (Ellis et al. 1996), open triangle represents the value from Stromlo-APM (Loveday et al. 1992), and the open diamond shows the ESP value (Zucca et al. 1997). Crosses show the estimates of Sawicki et al. (1997). In this paper we did not try to correct for the reddening effect of dust.

Finally, we must note that the above discussions do not account for the fact that the sample is selected at the I band, and the selection criterion is different for each redshift range. At $z < 0.5$, the sample is safely regarded as I-selected, while at $z > 3.0$, they are in fact rest-UV–selected. Thus, the ideal discussion of the evolution of the galaxy LF should be based on the suitably designed survey, as performed by Cowie et al. (1999). We will consider this point and take up more sophisticated discussions elsewhere (T. Takeuchi, in preparation).

5. SUMMARY AND CONCLUSION

The estimation of the LF from observational data is not a trivial task, because of the flux-limited nature of the astronomical data. We focus on the following four estimators: (1) the Schmidt-Eales ($1/V_{\rm max}$) method, (2) the Efstathiou–Ellis–Peterson (EEP) method, (3) the Chołoniewski method, and (4) the Lynden-Bell–Chołoniewski–Caditz–Petrosian (LCCP) method. We improved some of the estimators for studying the very distant universe, and examined their performances for a much wider class of functional forms by Monte Carlo simulation. We tested these four estimators by numerical experiments with mock catalogs. We also used the mock 2dF catalogs prepared by Cole et al. (1998). We then applied these estimators to the HDF photometric redshift catalog of Fernández-Soto et al. (1999). Our conclusions are as follows:

1. If the sample is spatially homogeneous, all estimators give consistent results with each other, and we did not find any bias for any LF shapes. Thus, when we have a sufficiently large galaxy sample, we can use any of the estimators examined in this paper. Even when the sample size is smaller, the mean values remain unbiased, although the standard deviations become larger.

2. Large fluctuations appear at the faint end of the LF, because the amount of available data is small. Therefore, the flatter the LF slope is, the larger the fluctuations become. When the sample size is small, fluctuations in the Chołoniewski method become seriously large due to shot noise, and thus we recommend this method for the analysis of large samples.

3. When a large cluster or void exists, the $1/V_{\rm max}$ estimator is severely affected in its LF shape estimation. The other three estimators are not at all affected by a cluster or void. They gave consistent results with each other, and the estimates showed perfect agreement with the input LFs.

4. We examined the calculation time of each method. Because of its algorithmic simplicity, the Chołoniewski method is the fastest among the four methods. The EEP method needs more iterations in the procedure than the others do, and a longer calculation time. The $1/V_{\rm max}$ method calculates the maximum volume, $V_{\rm max}$, for each galaxy, and

also needs significant calculation time. The LCCP method requires a large stack for the data-sorting procedure, which is a requirement of this method. Thus, we stress that the Chołoniewski method is the most economical from the standpoint of practical computing.

5. We examined more realistic large mock samples, specifically mock 2dF catalogs prepared by Cole et al. (1998). We found that the redshift distortion does not affect the LF estimates. When we treat such a large catalog, the advantage of the Chołoniewski method is extremely significant in terms of the computation time.

6. We derived the *I*- and *B*-band luminosity function of the HDF. The four different LF estimators gave consistent results for the HDF sample. We found the overall brightening of the LF. It seems that the faint end steepens toward $z = 2$–3, and settles down to the local value at $z \sim 3$. We note that the "fuzzy" redshift determination is known to affect the shape estimation (Liu et al. 1998).

7. We found a rather mild evolution of the LF. Despite the fact that the local LF is difficult to derive from the HDF data, our low-z value is consistent with other previous results. Our $\rho_L(B)$ at $0.5 < z < 1.0$ is also consistent with the CFRS highest redshift point within the errors. As a whole, our result is roughly consistent with that of SLY97, but it is lower at $2.0 < z < 3.0$. At very high z, we derived a moderately high ρ_L, implying that a significant number of stars have already formed at such a high redshift. We found that the evolution of $\rho_L(I)$ is flat.

First we would like to thank the anonymous referee for his useful suggestions and comments, which improved our paper very much in its clarity and English presentation. We offer our gratitude to Hiroyuki Hirashita and Fumiko Eizawa who gave a lot of useful suggestions. We also thank Kouji Ohta, Kouichiro Nakanishi, Toru Yamada, Takashi Ichikawa, and Kazuhiro Shimasaku for their fruitful discussions and useful comments. Mamoru Saitō, Hiroki Kurokawa, and Yasushi Suto are thanked for their continuous encouragements. This work owes a great debt to the photometric redshift catalog prepared by Fernández-Soto et al. T. T. T. and K. Y. acknowledge the Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists. We carried out the numerical computations and extensively used the databases at the Astronomical Data Analysis Center of the National Astronomical Observatory, Japan, which is an interuniversity research institute of astronomy operated by Ministry of Education, Science, Culture, and Sports.

APPENDIX

AKAIKE'S INFORMATION CRITERION

In this Appendix, we offer an informal introduction to Akaike's theory. The meaning of the maximum-likelihood method can be clearly understood by using the concepts of information theory. Since the middle of 1970s, vast advances have been made in the field of statistical inference through the discovery of Akaike's Information Criterion (AIC; Akaike 1974). The AIC is closely related to information entropy, especially the "relative entropy" of two probability distributions. The relative entropy has a property just like a distance in differential geometry, i.e., it is a distance between the two probability distributions. Using AIC enables us to compare the goodness of a certain model with that of another type directly. For this fascinating property, AIC is applied to various fields of studies. The AIC is expressed as

$$\mathrm{AIC} = -2[\ln \mathscr{L}(\hat{\theta}) - K] ,$$

where $\mathscr{L}$ is a likelihood function, $\hat{\theta}$ is a set of maximum-likelihood estimators, and K is the number of free parameters of the assumed model. The "most preferred" model is the one that minimizes the AIC.

Here we present the problem of polynomial regression model selection by using AIC as an example. As we mentioned in § 4, we adopt this procedure to determine the order of the K-correction as a function of redshift. Given a set of n pairs of observations, $(x_1, y_1), \ldots, (x_n, y_n)$, we fit the mth order polynomial model

$$y_i = \sum_{\ell=0}^{m} a_\ell x_i^\ell + \epsilon_i , \tag{A1}$$

where ϵ_i is an independent random variable that follows the normal distribution with mean 0 and dispersion σ^2. The variables x_i and y_i are called the explanatory variable and the objective variable, respectively. This model is a conditional distribution, of which the distribution of the objective variable y is a normal distribution, $f(y_i)$, with mean $a_0 + a_1 x_i + \cdots + a_m x_i^m$ and variance σ^2, i.e.,

$$f(y_i \,|\, a_0, \ldots, a_m, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{1}{2\sigma^2}\left(y_i - \sum_{\ell=0}^{m} a_\ell x_i^\ell\right)^2\right]. \tag{A2}$$

Therefore, when a set of data is $(x_1, y_1), \ldots, (x_n, y_n)$, the likelihood is given by

$$\begin{aligned}\mathscr{L}(y_1, \ldots, y_n \,|\, a_0, \ldots, a_m, \sigma^2) &= \prod_{i=1}^{n} f(y_i \,|\, a_0, \ldots, a_m, \sigma^2) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n/2} \prod_{i=1}^{n} \exp\left[\frac{1}{2\sigma^2}(y_i - \sum_{\ell=0}^{m} a_\ell x_i^\ell)^2\right].\end{aligned} \tag{A3}$$

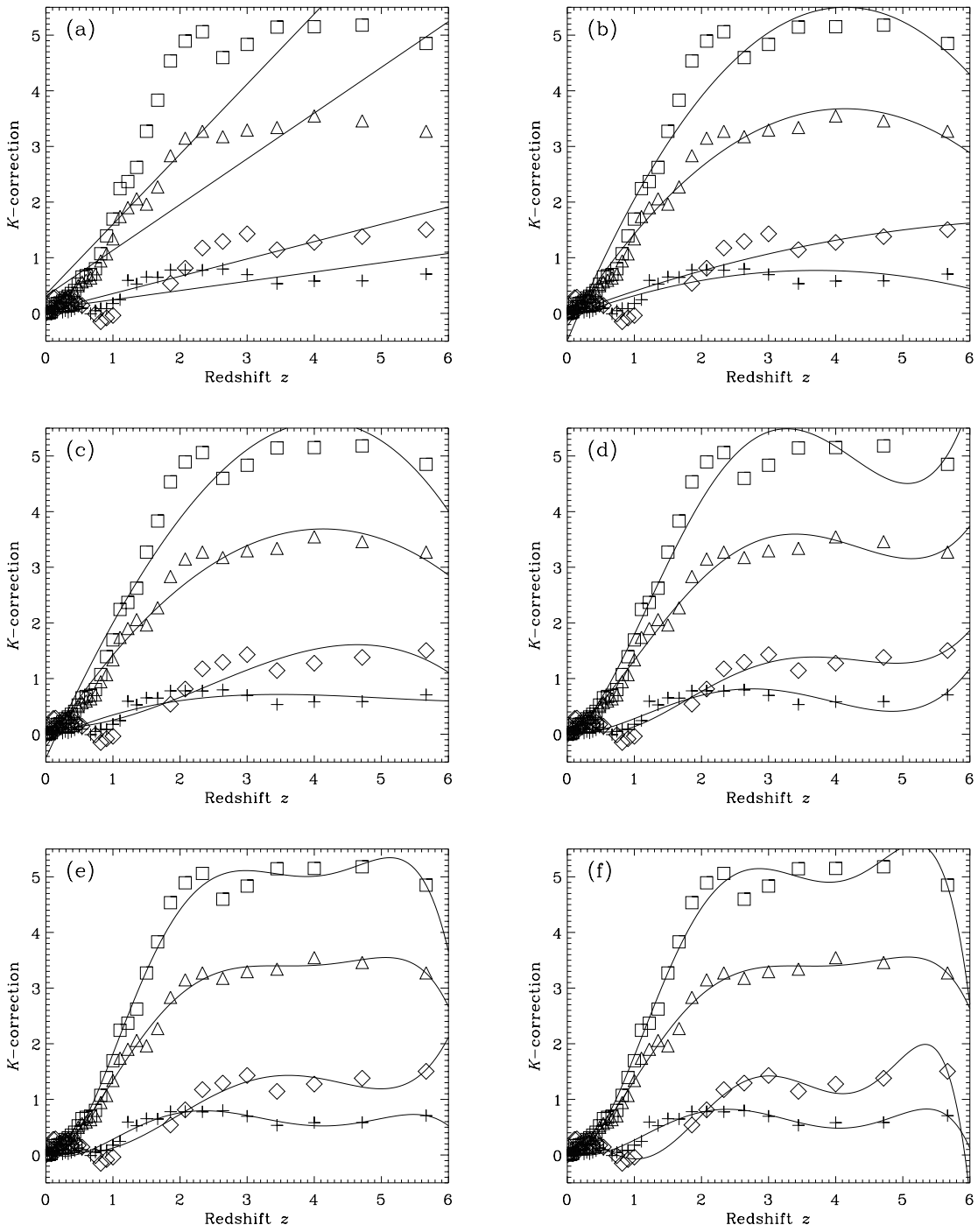

FIG. 24.—Result of the polynomial fitting to the K-correction of galaxies constructed from the data prepared by Kinney et al. (1996). Symbols represent the representative SEDs of galaxies: *open squares*: elliptical; *open triangles*: Sbc; *diamonds*: Scd; and *crosses*: Irr. Panels *a–f* correspond to the 1st through 6th fitting orders, respectively.

TABLE 1

AIC OF THE POLYNOMIAL FITTING MODEL FOR K-CORRECTION

	POLYNOMIAL FITTING ORDER					
TYPE	1st	2d	3d	4th	5th	6th
Elliptical.....	96.4724	41.6174	42.9716	19.0966	−6.93137	2.50965
Sbc.....	63.4433	−9.27677	−7.30190	−26.7273	−45.7942	−32.1221
Scd.....	5.04120	4.66447	−0.536911	−12.5160	−11.0097	−24.8293
Irr.....	−7.86433	−31.3419	−31.0330	−49.8901	−65.2359	−57.3888

The log likelihood is then expressed as

$$\ln \mathscr{L}(y \mid a_0, \ldots, a_m, \sigma^2) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(y_i - \sum_{\ell=0}^{m} a_\ell x_i^\ell \right)^2 . \tag{A4}$$

The log likelihood equation (A4) is maximized with respect to $a_0, \ldots, a_m$ when

$$S \equiv \sum_{i=1}^{n} (y_i - \sum_{\ell=0}^{m} a_\ell x_i^\ell)^2 \tag{A5}$$

is minimized. Thus, in the case of polynomial model fitting, the maximum-likelihood procedure is equivalent to the least-squares method. The necessary conditions that $a_0, \ldots, a_m$ maximize S are the normal equations of the least square,

$$\begin{aligned} \frac{\partial S}{\partial a_0} &= -2 \sum_{i=1}^{n} \left(y_i - \sum_{\ell=0}^{m} a_\ell x_i^\ell \right) = 0 , \\ \frac{\partial S}{\partial a_1} &= -2 \sum_{i=1}^{n} x_i \left(y_i - \sum_{\ell=0}^{m} a_\ell x_i^\ell \right) = 0 , \\ &\vdots \\ \frac{\partial S}{\partial a_m} &= -2 \sum_{i=1}^{n} x_i^m \left(y_i - \sum_{\ell=0}^{m} a_\ell x_i^\ell \right) = 0 , \end{aligned} \tag{A6}$$

and the maximum-likelihood estimates $\hat{a}_0, \ldots, \hat{a}_m$ are obtained by solving these linear equations. In addition, the necessary condition that σ^2 maximizes equation (A4) is

$$\frac{\partial \ln \mathscr{L}}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^{n} \left(y_i - \sum_{\ell=0}^{m} \widehat{a_\ell} x_i^\ell \right)^2 = 0 . \tag{A7}$$

The maximum-likelihood estimate of the residual variance, σ^2, is

$$\widehat{\sigma^2} = -\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \sum_{\ell=0}^{m} \widehat{a_\ell} x_i^\ell \right)^2 = -\frac{1}{n} \left(\sum_{i=1}^{n} y_i^2 - \sum_{\ell=0}^{m} \widehat{a_\ell} \sum_{i=1}^{n} x_i^\ell y_i \right) . \tag{A8}$$

Hereafter we denote the residual variance, σ^2, for a model with mth order as $\sigma^2(m)$. Then, from equations (A4) and (A8), the maximum log likelihood becomes

$$\ln \mathscr{L}(y \mid \widehat{a_0}, \ldots, \widehat{a_m}, \widehat{\sigma^2}) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \widehat{\sigma^2}(m) - \frac{n}{2} . \tag{A9}$$

The mth order polynomial model has $m + 2$ parameters $[a_0, \ldots, a_m, \sigma^2(m)]$. Substituting $K = m + 2$ and equation (A9) into equation (A1) gives the AIC of the mth order model,

$$\mathrm{AIC}(m) = n(\ln 2\pi + 1) + n \ln \widehat{\sigma^2}(m) + 2(m + 2) . \tag{A10}$$

The result of our polynomial fitting to the K-correction is presented in Figure 24. We also summarize the AIC value for each polynomial order in Table 1. The AIC values of elliptical, Sbc, and Irr in Table 1 really took their minima in the case that the fitting polynomials were those with 5th order, and only Scd data preferred the 6th order. Putting all accounts together, we chose the 5th order polynomial model.

REFERENCES

Akaike, H. 1974, IEEE Trans. Autom. Contrib., 19, 716
Avni, Y., & Bahcall, J. N. 1980, ApJ, 235, 694
Babu, G. J., & Feigelson, E. D. 1996, Astrostatistics (New York: Chapman & Hall)
Babul, A., & Ferguson, H. C. 1996, ApJ, 458, 100
Babul, A., & Rees, M. J. 1992, MNRAS, 255, 346
Baugh, G. M., Cole, S., & Frenk, C. S. 1996, MNRAS, 283, 1361
Beers, T. C. 1992, in Statistical Challenges in Modern Astronomy, ed. E. D. Feigelson & G. J. Babu (New York: Springer), 111
Bershady, M. A., Majewski, S. R., Koo, D. C., Kron, R. G., & Munn, J. A. 1997, ApJ, 490, L41
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bingelli, B., Sandage, A., & Tammann, G. A. 1988, ARA&A, 26, 509
Box, G. E. P., & Muller, M. E. 1958, Ann. Math. Statistics, 29, 610
Bruzual, G., & Charlot, A. 1993, ApJ, 405, 538
Caditz, D., & Petrosian, V. 1993, ApJ, 416, 450 (CP93)
Carswell, R. F. 1973, MNRAS, 162, 61
Chołoniewski, J. 1986, MNRAS, 223, 1 (C86)
———. 1987, MNRAS, 226, 273 (C87)
Cole, S., Hatton, S., Weinberg, D. H., & Frenck, C. S. 1998, MNRAS, 300, 945
Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393
Cowie, L. L., Songaila, A., & Barger, A. J. 1999, AJ, 118, 603
Cowie, L. L., Songaila, A., & Hu, E. M. 1996, AJ, 112, 839
Davis, M., & Huchra, J. P. 1982, ApJ, 254, 437
Davison, A. C., & Hinkley, D. V. 1997, Bootstrap Methods and Their Application (New York: Cambridge Univ. Press)
de Lapparent, V., Geller, M. J., & Huchra, J. P. 1989, ApJ, 343, 1
Eales, S. 1993, ApJ, 404, 51
Efron, B., & Tibshirani, R. 1993, An Introduction to the Bootstrap (New York: Chapman & Hall)
Efstathiou, G. 1996, in Cosmology and Large-Scale Structure, Les Houches Session LX, ed. R. Schaeffer et al. (Amsterdam: North-Holland), 133
Efstathiou, G., Ellis, R. S., & Peterson, B. A. 1988, MNRAS, 232, 431 (EEP)
Ellis, R. S. 1997, ARA&A, 35, 389
Ellis, R. S., Colless, M., Broadhurst, T. J., Heyl, J., & Glazebrook, K. 1996, MNRAS, 280, 235
Epanechnikov, V. A. 1969, Theor. Probab. Appl., 14, 153
Feigelson, E. D. 1992, in Statistical Challenges in Modern Astronomy, ed. E. D. Feigelson & G. J. Babu (New York: Springer), 221
Feigelson, E. D., & Nelson, P. I. 1985, ApJ, 293, 192
Felten, J. E. 1976, ApJ, 207, 700
———. 1977, AJ, 82, 861

Ferguson, H. 1998, in The Hubble Deep Field, ed. M. Livio, S. M. Fall, & P. Madau (Cambridge: Cambridge Univ. Press), 181
Fernández-Soto, A., Lanzetta, K. M., & Yahil, A. 1999, ApJ, 513, 34 (FLY99)
Furusawa, H., Shimasaku, K., Doi, M., & Okamura, S. 2000, ApJ, 534, 624
Gwyn, S. D. J., & Hartwick, F. D. A. 1996, ApJ, 468, L77
Heyl, J., Colless, M., Ellis, R. S., & Broadhurst, T. J. 1997, MNRAS, 285, 613
Hogg, D. W., & Phinney, E. S. 1997, ApJ, 488, L95
Hubble, E. 1936, The Realm of the Nebulae (London: Oxford Univ. Press)
Jackson, J. C. 1974, MNRAS, 166, 281
Kinney, A. L., Bohlin, R. C., Calzetti, D., Panagia, N., & Wyse, R. F. G. 1993, ApJS, 86, 5
Kinney, A. L., Calzetti, D., Bohlin, R. C., McQuade, K., Storchi-Bergmann, T., & Schmitt, H. R. 1996, ApJ, 467, 38
Kirshner, R. P., Oemler, A., Jr., & Schechter, P. L. 1979, AJ, 84, 951
Kitayama, T., & Suto, Y. 1997, ApJ, 490, 557
Kleinbaum, D. G. 1996, Survival Analysis (New York: Springer)
Knuth, D. E. 1998, The Art of Computer Programming, Vol. 2: Seminumerical Algorithms (3d Ed; Boston: Addison-Wesley)
Koo, D. C., & Kron, R. G. 1992, ARA&A, 30, 613
Koranyi, D. M., & Strauss, M. A. 1997, ApJ, 477, 36
Lehmann, E. L. 1999, Elements of Large-Sample Theory (New York: Springer)
Lilly, S. J., Tresse, L., Hammer, F., Crampton, D., & Le Fèvre, O. 1995, ApJ, 455, 108
Lin, H., Kirshner, R. P., Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., & Schechter, P. L. 1996, ApJ, 464, 60
Lin, H., Yee, H. K. C., Carlberg, R. G., Morris, S. L., Sawicki, M., Patton, D. R., Wirth, G., & Shepherd, C. W. 1999, ApJ, 518, 533
Liu, C. T., Green, R. F., Hall, P. B., & Osmer, P. S. 1998, AJ, 116, 1082
Loveday, J., Peterson, B. A., Efstathiou, G., & Maddox, S. J. 1992, ApJ, 390, 338
Lynden-Bell, D. 1971, MNRAS, 155, 95
Madau, P., Ferguson, H. C., Dickinson, M. E., Giavalisco, M., Steidel, C. C., & Fruchter, A. 1996, MNRAS, 283, 1388
Maloney, A., & Petrosian, V. 1999, ApJ, 518, 32
Marinoni, C., Monaco, P., Giuricin, G., & Constantini, B. 1999, ApJ, 521, 50
Marshall, H. L., Avni, Y., Tananbaum, H., & Zamorani, G. 1983, ApJ, 269, 35
Marzke, R. O., da Costa, L. N., Pellgrini, P. S., Willmer, N. A., & Geller, M. J. 1998, ApJ, 503, 617
Matsumoto, M., & Nishimura, T. 1998, ACM Trans. Modeling & Comput. Simul., 8, 3
Metcalfe, N., Ratcliffe, A., Shanks, T., & Fong, R. 1998, MNRAS, 294, 147
Mobasher, B., Rowan-Robinson, M., Georgakakis, A., & Eaton, N. 1996, MNRAS, 282, L7
Mobasher, B., Sharples, R. M., & Ellis, R. S. 1993, MNRAS, 263, 560
Nicoll, J. F., & Segal, I. E. 1983, A&A, 118, 180
Pascarelle, S. M., Lanzetta, K. M., & Fernández-Soto, A. 1998a, ApJ, 508, L1
Pascarelle, S. M., Windhorst, R. A., & Keel, W. C. 1998b, AJ, 116, 2659
Petrosian, V. 1992, in Statistical Challenges in Modern Astronomy, ed. E. D. Feigelson & G. J. Babu (New York: Springer), 173
Qin, Y.-P., & Xie, G.-Z. 1999, A&A, 341, 693
Sakamoto, Y., Ishiguro, M., & Kitagawa, G. 1986, Akaike Information Criterion Statistics (Dordrecht: Reidel)
Sandage, A., Tammann, G. A., & Yahil, A. 1979, ApJ, 232, 352 (STY79)
Sawicki, M. J., Lin, H., & Yee, H. K. C. 1997, AJ, 113, 1 (SLY97)
Schechter, P. L. 1976, ApJ, 203, 297
Schmidt, M. 1968, ApJ, 151, 393
Silverman, B. W. 1986, Density Estimation for Statistics and Data Analysis (London: Chapman & Hall)
Springel, V., & White, S. D. M. 1998, MNRAS, 298, 143
Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., Pettini, M., & Kellogg, M. 1998, ApJ, 492, 428
Strauss, M. A., & Willick, J. A. 1995, Phys. Rep., 261, 271
Stuart, A., Ord, K., & Arnold, S. 1999, Kendall's Advanced Theory of Statistics, Vol. 2A: Classical Inference and the Linear Model (6th Ed; London: Arnold)
Sturges, H. A. 1926, J. Am. Stat. Assoc., 21, 65
Subba Rao, M. U., Connolly, A. J., Szalay, A. S., & Koo, D. C. 1996, AJ, 112, 929
Szokoly, G. P., Subba Rao, M. U., Connolly, A. J., & Mobasher, B. 1998, ApJ, 492, 452
Takeuchi, T. T. 1999, Ap&SS, in press (preprint astro-ph/9909324)
Turner, E. L. 1979, ApJ, 231, 645
van Es, A. J. 1991, Aspects of Nonparametric Density Estimation (CWI tract 77; Amsterdam: Centrum voor Wiskunde en Informatica)
Viana, P. T. P., & Liddle, A. R. 1996, MNRAS, 281, 323
Weedman, D. W., Wolovitz, J. B., Bershady, M. A., & Schneider, D. P. 1998, AJ, 116, 1643
Williams, R. E., et al. 1996, AJ, 112, 1335
Willmer, C. N. A. 1997, AJ, 114, 898 (W97)
Zucca, E., et al. 1997, A&A, 326, 477