Astron. Astrophys. 356, 903–912 (2000)

ASTRONOMY
AND
ASTROPHYSICS

Statistics of low-mass companions to stars: Implications for their origin

T.F. Stepinski and D.C. Black

Lunar and Planetary Institute, 3600 Bay Area Blvd., Houston, TX 77058, USA (tom@lpis54.jsc.nasa.gov)

Received 31 August 1999 / Accepted 17 February 2000

Abstract. One of the more significant results from observational astronomy over the past few years has been the detection, primarily via radial velocity studies, of low-mass companions (LMCs) to solar-like stars. The commonly held interpretation of these is that the majority are "extrasolar planets" whereas the rest are brown dwarfs, the distinction made on the basis of apparent discontinuity in the distribution of $M \sin i$ for LMCs as revealed by a histogram. We report here results from statistical analysis of $M \sin i$, as well as of the orbital elements data for available LMCs, to test the assertion that the LMCs population is heterogeneous. The outcome is mixed. Solely on the basis of the distribution of $M \sin i$ a heterogeneous model is preferable, although no unique best-fit mixture can be determined. On the basis of the distribution of orbital periods and eccentricities a homogeneous model is strongly preferable. Overall, we find that a definitive statement asserting that LMCs population is heterogeneous is, at present, unjustified. In addition we compare statistics of LMCs with a compatible sample of stellar binaries. We find a remarkable statistical similarity between these two populations. This similarity coupled with marked populational dissimilarity between LMCs and acknowledged planets motivates us to suggest a common origin hypothesis for LMCs and stellar binaries as an alternative to the prevailing interpretation. We discuss merits of such a hypothesis and indicate a possible scenario for the formation of LMCs.

Key words: stars: binaries: spectroscopic – stars: formation – stars: low-mass, brown dwarfs – stars: planetary systems – stars: statistics

1. Introduction

The rising accuracy of radial velocity techniques has resulted in detection of numerous unresolved low-mass companions to solar type stars. So far (as of early 1999), surveys revealed 17 objects (Marcy et al. 1999) for which a projected mass, $M \sin i$, of a companion is smaller than 12 $M_{\rm J}$. The angle i is that between an observer's line-of-sight to a star and the normal to the orbital plane of the companion/star system. Because of their small projected masses, and thus expected small actual masses assuming random orientation of orbital planes in space, these companions have been classified as extrasolar planets (EP) or planet candidates (Marcy & Butler 1998; Marcy et al. 1999). In addition, 10 objects with $17 M_{\rm J} < M \sin i < 70 M_{\rm J}$ have been found (Mayor et al. 1997) and classified as brown dwarf candidates (BD), again on the basis of their expected masses being sub-stellar but substantially higher than the mass of Jupiter. Although this dual classification of low-mass companions (LMCs) is based solely on the magnitude of their projected masses, it has been also widely assumed (see the aforementioned references) that it reflects differences in origin. Specifically, it has been assumed that EPs formed via the process essentially identical to what is postulated for the formation of Jupiter in the Solar System – buildup by aggregation from a protoplanetary disk, whereas BDs formed presumably via cloud fragmentation, just like the stars. In this paper we address two distinct, yet interconnected issues. First, the mass distribution of LMCs and whether their division into EPs and BDs is statistically justified. Second, the statistics of LMC's orbital elements and what they may imply regarding the origin of LMCs.

We start by enumerating the principal arguments advanced for dividing LMCs into EP and BD:

Mass distribution of LMCs. This argument as generally presented stems from a histogram of $M \sin i$ from all available LMC data (Marcy & Butler 1998; Mazeh et al. 1998; Marcy et al. 1999). Such a histogram shows a spike in the first bin containing LMC with the smallest masses followed by subsequent bins containing very few objects (see Fig. 2). Because histograms are supposed to portray the underlaying probability distribution function (PDF), the proponents of the dual character of LMCs argue that the actual PDF of $M \sin i$ is *discontinuous* at some small value of $M \sin i$ providing a natural divide and observationally defining two sub-populations of LMC.

However, in cases where a number of objects in the sample is relatively small and there is reason to believe that the underlying PDF is skewed, histograms are poor indicators of an actual PDF. In this paper we infer the *functional form* of the PDF from the empirical cumulative distribution function (CDF), and determine the *parameters* of the PDF using maximum-likelihood estimation (MLE).

Send offprint requests to: T. Stepinski

Minimum mass of brown dwarf. This argument stems from the alleged lower limit to the mass of brown dwarfs based upon the concept of opacity-limited fragmentation. Estimates of this mass limit yield values of about 10 $M_{\rm J}$ (Rees 1976; Silk 1977), although lower estimates are possible if fragmentation occurs in a disk (Boss 1998). Such a limit could provide theoretical support to the notion of duality of the LMC population, provided it falls near the purported mass cutoff. However, this limit must be considered as highly uncertain quantitatively . Additionally, the "minimum mass" argument ignores the possibility that evolutionary effects such as mass exchange have altered the masses of the closer companions. The strength of this argument as support for a heterogeneous LMC population is marginal at best.

Mass-eccentricity relation. This argument arises from plotting LMC projected masses versus their eccentricities. Supporters of the dual character of LMCs have pointed out that LMC below a certain mass have low eccentricities, and those above that mass have high eccentricities, again revealing a "discontinuity" that suggests the existence of two sub-populations (Mayor et al. 1998). This argument seems to be a historical footnote as new data do not conform to the alleged mass-eccentricity relation.

Metallicity. The fourth argument given for a dual character of LMCs comes from the metallicities of stars with LMCs. Stars with LMC designated as EP are metal rich compared to field stars (Gonzalez 1998; Gonzalez et al. 1999). However, as no direct comparison of metallicities between parent stars of designated EP and parent stars of designated BD has been published, the metallicity argument does not at present contribute to the question of homogeneity or heterogeneity of the LMC population.

The case for the existence of two distinct species in the population of LMC is deserving of a more extensive treatment than it has received in the literature to date. In this paper we concentrate on evaluating two of the above arguments using statistical analysis of data relating to all 27 LMC. Our goal is to estimate from available data the underlying PDFs for projected masses, periods, and eccentricities of LMC.

Our approach will be to employ a *parametric* statistical model in which the data is assumed to be drawn from a mixture of two PDFs (one describing putative EP and the second describing putative BD) each having a specific form (inferred from the empirical CDF), but undetermined parameters. In addition, the parameter describing the relative contribution of two components to the overall mixture is undetermined. MLE is used to determine all unknown parameters. This approach distinguishes our work from that of Heacox (1999) who employed a nonparametric statistical model to analyze distributions of various LMC quantities.

Sect. 2 discusses data adjustments and Sect. 3 contains a description of our statistical analysis. In Sect. 4 we present results pertaining to projected masses. Separately, in Sect. 5 we present results pertaining to periods and eccentricities. Finally, in Sect. 6, we present conclusions and discussion.

2. Data adjustments

LMC data considered in this paper come from several different surveys. This fact puts in question the representativeness of the overall LMC sample and thus its suitability for statistical analysis. To alleviate this problem some adjustments are needed when joining LMC data from different surveys into a single set. In the context of low-mass and stellar-mass companions such adjustments are discussed by Mazeh et al. (1998). The corrections take into account the effects due to instrumental precision and number of stars examined in the various radial velocity surveys. In addition, Mazeh et al. (1998) correct for the $\sin i$ factor because they are interested in a distribution of an actual mass of companions rather than a distribution of a projected mass.

We collect our LMCs sample from numerous surveys, but it is only necessary to consider two distinct categories, objects obtained from relatively low precision ($\sim$ 300 m/sec) survey of 570 stars by Mayor et al. (1997) and objects obtained from relatively high precision ($\sim$ 10 m/sec) surveys of about 300 stars (see Marcy et al. 1999 and references therein).

A correction protocol described by Mazeh et al. (1998) is valid assuming that low and high precision surveys are compatible, statistically independent and unbiased. However, due to differences in target selection criteria, different surveys are not entirely compatible, and are likely not to be statistically independent. Therefore, it is not clear what adjustment protocol, if any, is the most appropriate. Given these uncertainties we use both unadjusted and adjusted LMCs data to infer the distribution of $M \sin i$. We do not correct the data for the $\sin i$ factor because we restrict ourselves to studying the distribution of *projected* mass only. This is dictated by the small size of LMCs sample. Thus, the names EP and BD have to be taken with caution inasmuch as $M \sin i$ is used as a surrogate for an actual mass. Finally, only unadjusted data is used to infer distributions of LMCs orbital parameters.

3. Statistical model

We look for the underlying PDFs for projected masses, periods, and eccentricities using the MLE. Such an estimation maximizes the probability of drawing a particular datum that was in fact obtained. This approach requires specifying the functional form of the PDF and estimating the values of free parameters. The form of the PDF can be deduced from the empirical CDF constructed for LMC quantities. Let $\bar{y} = (y_1, y_2, ..., y_N)$ be a list of either projected masses, periods, or eccentricities for N LMCs and assume that $\bar{y}$ has been already sorted by size in increasing order. Then the empirical CDF, denoted by $F(y)$ is defined by

$$F(y) = \begin{cases} 0, & y < y_1 \\ \frac{i}{N}, & y_i \leq y < y_{i+1} \\ 1, & y_N \leq y \end{cases} \quad (1)$$

The estimation process is complicated by the fact that we have to allow for the existence of two sub-populations in the overall population of LMC. We assume that $\bar{y}$ is drawn from a mixture of two PDFs, $f_1(y|\theta_1)$ which describes the distribution of quantity y for "planets," and $f_2(y|\theta_2)$ which describes the distribution of quantity y for "brown dwarfs," where θ_1 and θ_2 are lists of parameters characterizing respective PDFs. Thus, the PDF for the entire LMC population can be expressed as follow,

$$f(y|\theta_1, \theta_2, \varepsilon) = (1 - \varepsilon) f_1(y|\theta_1) + \varepsilon f_2(y|\theta_2) \quad (2)$$

where $0 < \varepsilon < 1$ is a mixture parameter. Drawing (observing), say projected masses $M \sin i$ from a total LMC population distributed according to (2) can be interpreted as a two step process. First a Bernoulli random variable b is drawn taking a value of 1 with probability $(1 - \varepsilon)$, or value 2 with probability ε. According to the value of b, $M \sin i$ is then drawn from one of the two sub-populations with PDFs $f_1(y|\theta_1)$ and $f_2(y|\theta_2)$. We assume that the "allocation" variable b is not directly observed. This means that we don't a priori put any labels on the data. The labels, if any, can be put a posteriori if indicated by statistical analysis. The complete data is thus $\bar{z} = (z_1, z_2, ..., z_N)$, where $z_j = (y_j, b_j)$. The PDF given by (2) can be interpreted as $g(z|\theta)$ with $\theta = (\theta_1, \theta_2, \varepsilon)$. The log-likelihood function formed from the data is

$$\log L = \log \prod_{j=1}^{N} g(z_j|\theta) = \sum_{j=1}^{N} \log g(z_j|\theta) \quad (3)$$

A MLE is a value of θ denoted by $\hat{\theta}$ that maximizes $\log L$. In general, obtaining $\hat{\theta}$ is a nontrivial undertaking because θ is a vector of potentially many dimensions and $g(z|\theta)$ can be a complicated function. We use the Expectation-Maximization (EM) numerical algorithm (Dempster et al. 1977) to find a MLE. Note that this estimate contains the mixture parameter ε. If the estimation of ε is close to zero, a homogeneous population is indicated.

4. Projected masses

We carried out calculations for several cases set apart by different adjustments to the LMC data, no adjustments, adjustment for sample size, and adjustments for both sample size and precision. Adjustments are achieved by enlarging the population of objects in a certain projected mass range by an appropriate factor. To correct for sample size we enlarged the population of EP by the factor of 2. Following Mazeh et al. (1998) we correct for instrumental precision by further enlarging the population of planets with $M \sin i < 1 M_{\rm J}$ and BD with $10 M_{\rm J} < M \sin i < 30 M_{\rm J}$ by another factor of 2. It should be noted that Mazeh et al. use a 2σ (where σ is a measurement error) criterion for establishing the minimum detectable EP signal (4σ peak-to-peak). This is in contrast to the 4σ semi-amplitude criterion suggested by Marcy & Butler (1998) and used by us later in this paper. Use of this more stringent detection criterion would yield a modest increase in the correction factor for the low end of the EP data set, but it would not alter the conclusion.

The first step is to calculate an empirical CDF for LMC projected masses. Displaying a CDF on the log-linear scale makes an identification of the underlying PDF easier. In such a scaling any straight line corresponds to a PDF having a power-law form, $f \sim y^{-p}$, with the index $p = 1$. Convex departures from the straight line indicate PDF in the form of the power-law with $p < 1$, whereas concave departures from the straight line indicate power-law PDF with $p > 1$. Similarly, PDFs in other forms (for example, normal distribution, log-normal distribution etc.) have their own characteristic CDF signatures. In the case when the gradient of the empirical CDF changes abruptly, a mixture of two PDFs is indicated.

The empirical CDF for projected masses of LMCs (regardless of considered adjustments) can be best characterized as either a single smooth curve quite close to a straight line, or a piecewise-smooth curve with two component curves. Thus, we infer from the data that the PDF of projected masses has a functional form that is either a single power-law, or a mixture of two power-laws. Of course the empirical CDF constructed from only 27 data points cannot be used to unambiguously infer the underlying PDF and it is conceivable that the data came from a distribution having functional form different from what we have inferred. Newertheless, a power-law provides the least structured candidate for the underlying PDF. Therefore we adopt the following form for the PDF of LMC projected masses

$$\begin{aligned} f(M \sin i|\theta) \; = \; & (1 - \varepsilon) A_1 (M \sin i)^{-p_1} H_1 \\ & + \varepsilon A_2 (M \sin i)^{-p_2} H_2 \end{aligned} \quad (4)$$

where H_1 and H_2 are cut-offs defined in terms of the Heaviside step function H,

$$\begin{aligned} H_1 &= H[M \sin i - (M \sin i)^{\rm min}_{\rm ep}]\, H[(M \sin i)^{\rm max}_{\rm ep} - M \sin i] \\ H_2 &= H[M \sin i - (M \sin i)^{\rm min}_{\rm bd}]\, H[(M \sin i)^{max}_{\rm bd} - M \sin i], \end{aligned} \quad (5)$$

In other words, the PDF consists of two components, the EP component which is a power-law with the index p_1 and valid for projected masses between $(M \sin i)^{\rm min}_{\rm ep}$ and $(M \sin i)^{\rm max}_{\rm ep}$, and the BD component given by a power-law with the index p_2 and valid for projected masses between $(M \sin i)^{\rm min}_{\rm bd}$ and $(M \sin i)^{\rm max}_{\rm bd}$. Values of $(M \sin i)^{\rm min}_{\rm ep} = 0.3 M_{\rm J}$ and $(M \sin i)^{\rm max}_{\rm bd} = 70 M_{\rm J}$ are set by observations, but there are no a priori assumptions about values of $(M \sin i)^{\rm max}_{\rm ep}$ and $(M \sin i)^{\rm min}_{\rm bd}$; the distributions may, in principle, overlap, connect, or there may be a gap between them. The parameter list θ has five components, p_1, p_2, $(M \sin i)^{\rm max}_{\rm ep}$, $(M \sin i)^{\rm min}_{\rm bd}$, and ε, because we decided to fix values of $(M \sin i)^{\rm min}_{\rm ep}$ and $(M \sin i)^{\rm max}_{\rm bd}$. Constants A_1 and A_2 are to assure that contributing PDF s integrate to 1. They are expressible in terms of already defined parameters.

Our goal is to determine the MLE of θ. We set up our calculations as follows. We allow both $(M \sin i)^{\rm max}_{\rm ep}$ and $(M \sin i)^{\rm min}_{\rm bd}$ to be any value from $5 M_{\rm J}$ to $50 M_{\rm J}$ in steps of $2.5 M_{\rm J}$. Thus, there are altogether $19^2 = 361$ possible PDFs under consideration. For each possible PDF with the pre-defined mass domain we employ the EM algorithm which finds the MLE of p_1, p_2, and ε and record the corresponding (maximized) value of $\log L$. Note that, in principle, the EM algorithm should be able to find the MLE of the entire θ, without auxiliary cycling through two

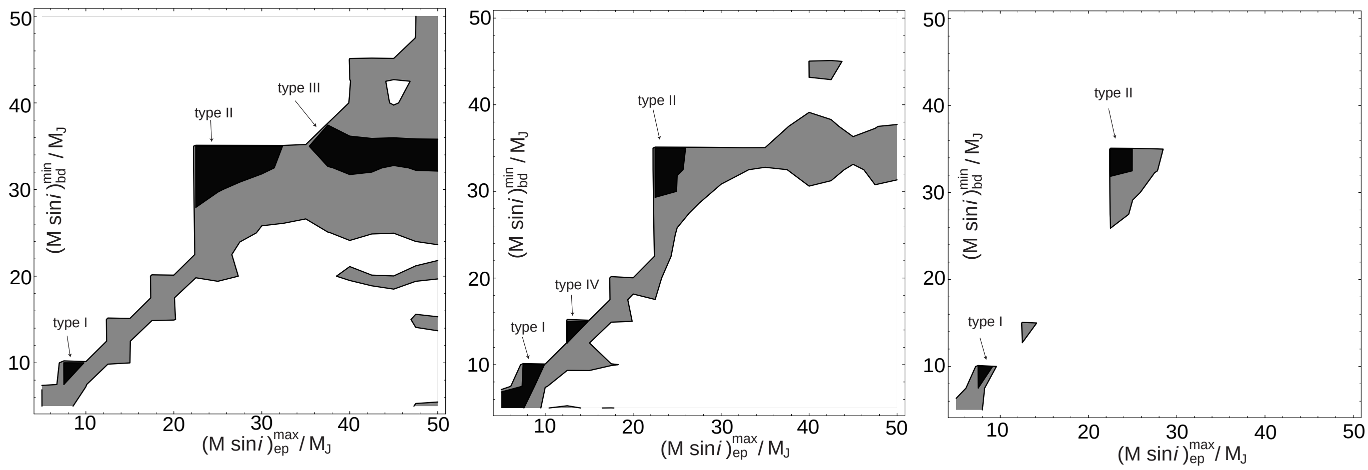

Fig. 1. The summary of testing the hypothesis that the PDF of $M \sin i$ for LMCs are given by a single power-law against the alternative hypothesis that it is given by a mixture of two power-laws. Possible mixture PDFs are indexed by $(M \sin i)^{\rm max}_{\rm ep}$ and $(M \sin i)^{\rm min}_{\rm bd}$). The single power-law hypothesis is accepted over the mixture hypothesis for mixtures in the white region. The mixture hypothesis is accepted at the significance level $s \leq 1$ for mixtures in the gray region. The black subset of the gray region contains mixtures accepted at $s \leq 0.05$. These, best fit models, can be grouped into several types as indicated by arrows. The panels from left to right are for the unadjusted LMC data, data adjusted for sample size, and data adjusted for both sample size and instrumental precision.

of its components. However, due to the special character of these parameters (they define cut-offs of PDFs) we find such a straightforward application of the EM algorithm difficult to implement. We also calculate the MLE of $\theta_0 = p$ and record the maximized value of $\log L$ for a PDF given by a single power-law, $f \sim (M \sin i)^{-p}$, over the entire domain of masses.

In the above context, the best way to address the main issue of heterogeneity versus homogeneity of the LMC population is to test the hypothesis that the PDF of $M \sin i$ is given by a single power-law against the alternative hypothesis that it is given by a mixture of two power-laws. The test is simple, we compare the value of $(\log L)^{\rm sg}$ for the MLE-determined single power-law PDF with values of $(\log L)^{\rm mix}_k$ for all $k = 1, \ldots, 361$ MLE-determined mixture PDFs. A mixture PDF hypothesis is a contender only for such mixtures (indexed by k) for which $(\log L)^{\rm mix}_k > (\log L)^{\rm sg}$. For all contending mixtures we calculate a significance level, $s = 1 - {\rm CDF}_{\chi^2}[-2((\log L)^{\rm sg} - (\log L)^{\rm mix}_k)]$, using the χ^2 distribution. The value of s is a probability that the single power-law hypothesis is falsely rejected. The level at which one accepts the heterogeneous hypothesis is, of course, subjective. Statistics textbooks (for example Mack 1967) give following guidline; a significance level of 0.05 is equated with "just significant" and a level of 0.01 with "highly significant."

Fig. 1 shows the results of the test. Each point on the graph corresponds to a mixture PDF indexed by $[(M \sin i)^{\rm max}_{\rm ep}, (M \sin i)^{\rm min}_{\rm bd}]$ instead of k. The white area correspond to mixture PDFs for which $(\log L)^{\rm mix} < (\log L)^{\rm sg}$ and the gray area indicate contending mixtures. The subset of contending mixtures for which $s \leq 0.05$ is colored black. Thus, the black regions indicate mixture models that should be accepted over the best homogeneous fit.

The first result is that, in all considered cases, our formal procedure locates some mixture models that are better fits than the best single power-law model. However, there is no *unique* best-fit mixture, instead, in all cases, the best mixture fits can be grouped into several types set apart by their overall character.

The best single power-law fit to the unadjusted LMC data has an index $p = 0.89$ (± 0.018). Three distinct mixture types yielding significantly better fit than the single power-law model can be identified. Type I is a mixture with a power-law break at $M \sin i$ about 5-10$M_{\rm J}$. The PDF for small projected masses (before the break) is steeper than after the break. Type I fits suggest a heterogeneous LMCs population with a character much like the one usually implied in the literature (for example Marcy et al. 1999). In a Type II mixture the power-law index is close to 1.0 over the entire range of LMCs projected masses but the EP PDF ends at $M \sin i \approx 22 M_{\rm J}$ and the BD PDF starts at $M \sin i \approx 35 M_{\rm J}$. Type II fits would be consistent with the findings of Mazeh et al. (1998). Finally, in the Type III fit the power-law index is also close to 1.0 over the entire range of projected masses, but the high end of the EP PDF overlaps with the low end of the BD PDF in the region of $M \sin i \approx 35 - 40 M_{\rm J}$.

The formally best fit to the unadjusted LMC data is of Type II with $s = 0.008$. Fig. 2 shows best fits for a single power-law model and for all types of preferred mixtures. CDFs are the best tools to visualize a fitness of a model to the data. We have, however, also constructed histograms based on corresponding models. Such histograms are constructed by integrating the model PDF between the cutpoints defining the bins and rounding results to the nearest integer. Comparison between data-derived and model-derived histograms offer an alternative way to visually judge the fitness of a model.

The best single power-law to the sample size adjusted LMC data has an index $p = 1.06$ (± 0.014). As was the case with the

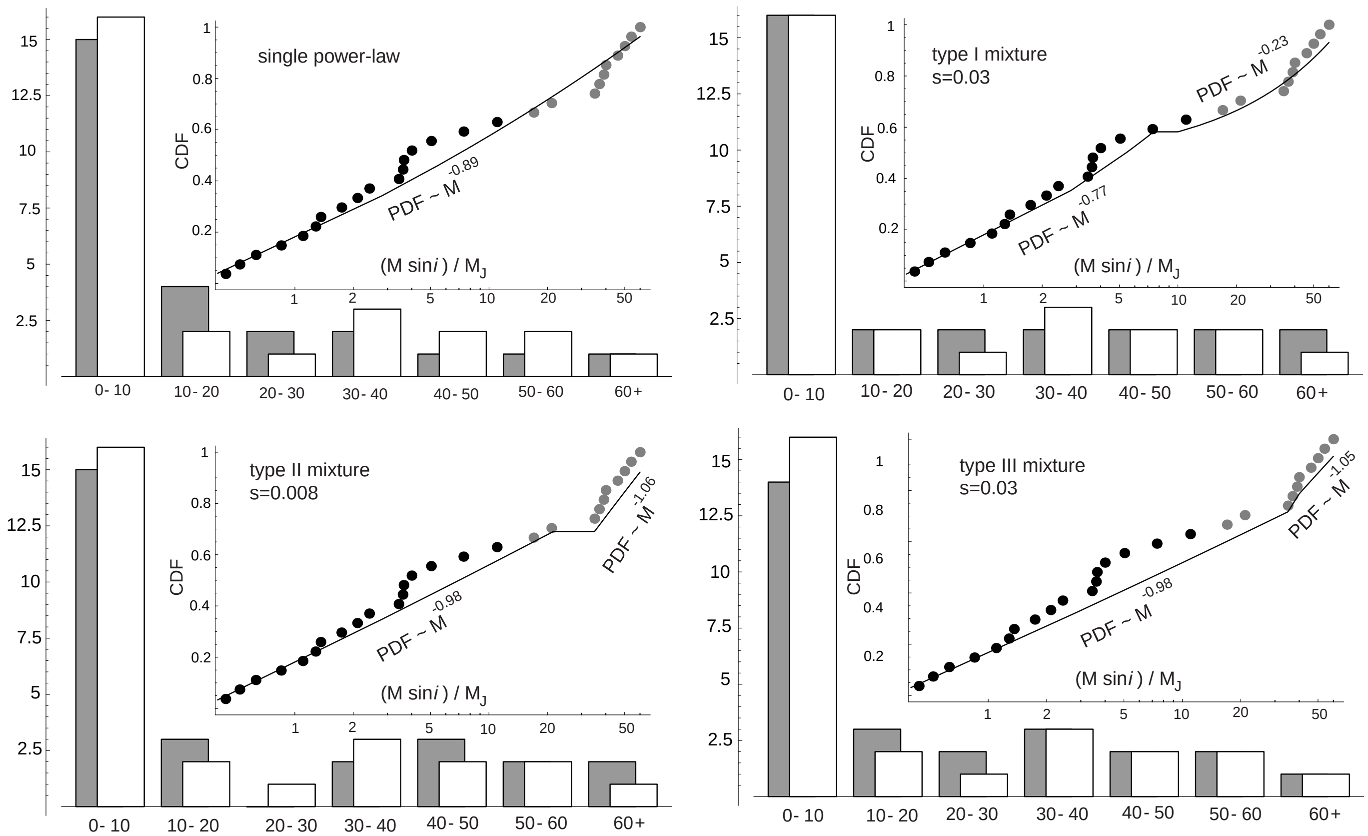

Fig. 2. Estimations of the PDF of $M \sin i$ using unadjusted LMC data. Four panels correspond to four different estimations. All panels show the empirical CDF for $M \sin i$, black dots indicate contribution from EP and gray dots contributions from BD. The histogram with white bars is produced from observations. Solid lines show CDFs corresponding to indicated MLE-determined PDFs. Upper left panel is for a single power-law PDF, other panels are for best mixture models as indicated by their type and significance level. The histogram with gray bars is produced from corresponding PDF.

unadjusted LMC data, Type I and II mixture fits are statistically superior to a homogeneous fit, but Type III fits are not. A new type of heterogeneous fit, Type IV, is present. Type IVs are similar to the Type I fits, but with the power-law break located at $M \sin i = 12 - 15 M_{\rm J}$. The formally best fit to the sample size adjusted LMC data is of Type I with $s = 0.004$.

Finally, the best single power-law to the sample size and precision adjusted LMC data has an index $p = 1.15$ (± 0.007). In such case only Type I and II mixture are better than the single power-law. The formally best fit to the size and instrument precision adjusted LMC data is of Type I with $s = 0.02$.

Overall, our calculations show that heterogeneous models can be found that fit the LMC data better than a homogeneous model. However, given presently available data we cannot pinpoint an unique mixture model. Two mixture types seem to offer comparable fits. One (Type I) suggests the LMCs population divided roughly at the theoretical lower limit to the mass of brown dwarfs based upon the concept of opacity-limited fragmentation. A second (Type II) suggests possible discontinuity at $M \sin i$ of about 20-35$M_{\rm J}$. This result is robust inasmuch as it is independent of possible data adjustments. Evidently more observations are needed to settle this issue. Note that single power-law models offer reasonable, although formally worse, fits. Fits to the adjusted LMC data sets yield steeper single power-law models, although, the range from $p = 0.89$ (± 0.018), for unadjusted data, to $p = 1.15$ (0.007), for the most adjusted data, is not dramatic.

Clear resolution of the nature of the projected mass distribution awaits the results from a comprehensive survey of a large number ($\sim$1000) stars using one instrument. Such a survey is just beginning in the Southern Hemisphere (Queloz, private communication).

5. Periods and eccentricities

Studying the projected mass distribution of LMCs is the way (apart from the $\sin i$ difficulty) of addressing the issue of their character that has been emphasized by other workers. However, examining distributions of the orbital periods and eccentricities of the LMCs may provide a clearer sense of the nature of these objects as these observables have no ambiguities associated with their value. Additionally, they are relatively, but not completely immune to the question of completeness as the entire range of values that these variables can take is accessible to all surveys. The exception to this lies in the longer periods, but that is not likely to influence the results that are discussed below. Toward

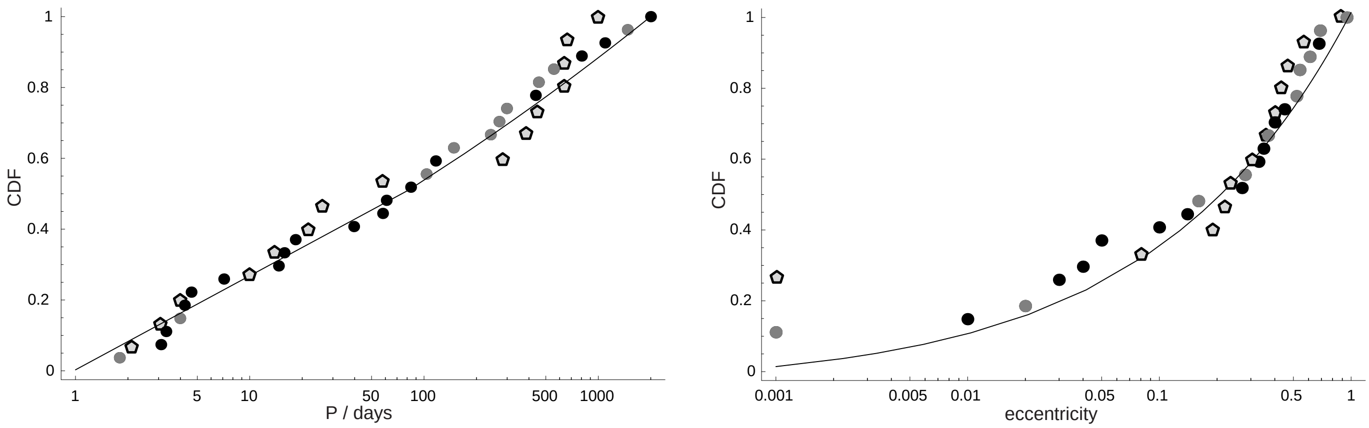

Fig. 3. Empirical CDFs of periods (left) and eccentricities (right) for LMC (denoted by dots) and selected stellar companions (denoted by pentagon symbols). Black dots indicate contributions from designated extrasolar planets and gray dots indicate contributions from designated brown dwarfs. Curves are model CDFs calculated from MLE-determined single power-laws PDFs. Because of log-linear scale objects with $e = 0$ are plotted as $e = 0.001$.

that end we have constructed empirical CDFs for periods and eccentricities of LMC.

Note that the character of the problem for periods and eccentricities is qualitatively different from that for the projected masses. In the case of the projected mass distribution, one could imagine a mixture PDF with components PDFs having mostly separated domains, with some possible overlap. However, in the case of periods and eccentricities, if actual PDFs are mixtures, their components overlap over the entire domain. Nevertheless, the EM algorithm can still be used to test homogeneity versus heterogeneity of the sample.

We use the same technique as described in Sect. 4 to test the homogeneity versus the heterogeneity of the population of LMCs with respect to distributions of their orbital periods and eccentricities. The empirical CDF for the orbital periods of LMCs, plotted on the log-linear scale can be best characterized as either a single straight line or a piecewise-smooth curve with two components. Thus we infer from tha data (see also Sect. 4) that the functional form of the PDF of LMC periods is a binary mixture (2) with f_1 and f_2 given by power-laws with different indices, but having the same domain consisting of the entire range of observed periods. We calculate the MLE of $\theta = (p_1, p_2, \varepsilon)$. Calculations reveal that the log-likelihood function (3) is minimized for $\varepsilon = 0$ and $p = p_1 = 0.98$ (± 0.01). Therefore, there is no evidence of two populations in the LMC in the available period data. Fig. 3 (the left panel) shows the empirical CDF for the orbital periods of all LMCs companions together with the MLE-estimated fit.

Also shown in Fig. 3 is the empirical CDF for orbital periods of selected stellar companions to solar-type stars. The particular selection of 15 binaries, due to Heacox (1999), is designed to be compatible with LMCs. It is a subset of the binaries in the Duquennoy & Mayor (1991) survey constrained by the requirement that primaries are population I and semi-major axes are less than 3 AU. Note that the empirical CDF for periods for stellar companions seems to be indistinguishable from that defined by the LMC. Formally, the MLE-estimated single power-law PDF for periods using the Heacox binaries data has $p = 0.89$ (± 0.05). We have also constructed the empirical CDF for orbital periods of secondaries using all 52 spectroscopic binaries from the survey by Duquennoy & Mayor (1991). Again, the shape of the empirical CDF suggest a single power-law, and the MLE-estimated power-law index is $p = 0.87$ (± 0.004).

The empirical CDF for eccentricities of LMCs, plotted on the log-linear scale can be best characterized as either a single convex curve or a piecewise-smooth curve with two components. Thus, we infer from the data (see also Sect. 4) that the functional form of the PDF of LMC eccentricities is a binary mixture of two power-laws with different indices and a common domain. The log-likelihood function is minimized for $\varepsilon = 0$. Therefore, as in the case of periods distribution, there is no evidence of two populations in the LMC in the available eccentricity data. The MLE-estimated single power-law PDF for eccentricities using all LMCs with $e > 0.001$ has $p = 0.64$ (± 0.03). Fig. 3 (the right panel) shows the empirical CDF for eccentricities of all LMCs companions together with the best power-law fit. Also shown in Fig. 3 is the CDF for eccentricities of secondaries in the aforementioned sample of stellar binaries. As in the case of orbital periods, the empirical CDFs for eccentricities for stellar companions and LMCs are very similar. Formally, the MLE-estimated single power-law PDF for eccentricities using the Heacox binaries with $e > 0.001$ has $p = 0.63$ (± 0.08). We have also constructed the empirical CDF for eccentricities of secondaries using all 52 spectroscopic binaries from the survey by Duquennoy & Mayor (1991). Single power-law is indicated, and the MLE-estimated power-law index is $p = 0.63$ (± 0.04) identical to that obtained for the Heacox subset. It appears that LMCs and stellar binaries have orbital elements distributed alike. We shell return to this key facet of the LMCs in the discussion section.

So far we have considered individual distributions of projected masses and orbital elements of LMC and stellar companions. However, we can also study dual correlations between these quantities. The clearest correlation is that between peri-

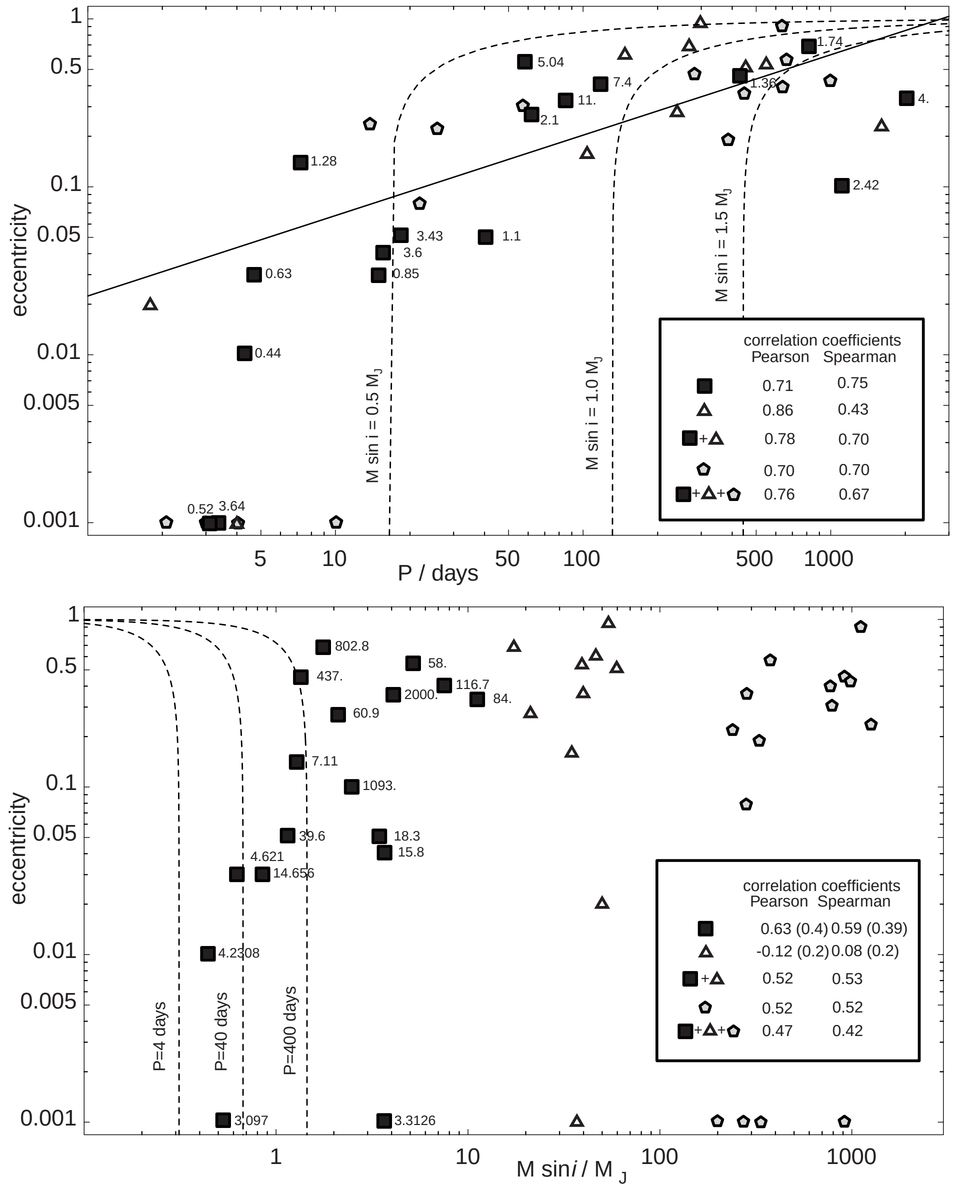

Fig. 4. $P-e$ (upper panel) and $M \sin i-e$ (lower panel) diagrams. Square symbols denote EP, triangle symbols denote BD, and pentagon symbols denote stellar companions. Because of logarithmic scale objects with $e = 0$ are plotted as $e = 0.001$. The EP on the upper panel are labeled by their projected mass (in units of $M_{\rm J}$) and the EP on the lower panel are labeled by their periods (given in days). The solid line on the $P-e$ diagram represents the least square fit using all data points except those with $e = 0.001$. Dashed curves denote selected detectability limits.

ods and eccentricities. Black (1997) has previously noted this. Fig. 4 (upper panel) shows the $P-e$ diagram composed of all LMCs and selected stellar companions. The insert lists correlation coefficients for various sub-groups of the data. The Spearman coefficient, $\rho_{\rm s}$, measures the correlation between rankings of periods and eccentricities, it gauges the strength of the associations between two variables. Perfect concordance of both rankings yields $\rho_{\rm s} = 1$ and indicates a direct *causal relation* between both quantities. Smaller values of $\rho_{\rm s}$ indicate an existence of a *trend* rather than a one-to-one relation, $\rho_{\rm s} = 0$ indicates no association. The fact that $\rho_{\rm s} \neq 0$ for all entries in the insert is statistically significant at the $s = 0.05$ level, except for the BD population, for which it is significant, but only at the $s = 0.25$ level. This has been determined using the fact that the quantity $\rho_{\rm s}\sqrt{(n-1)/(1-\rho_{\rm s}^2)}$ has a Student t-distribution with $n-2$ degrees of freedom. These correlations have been calculated from the data set excluding objects in orbits suspected of being altered by stellar tides, i.e., those with orbital periods of a few days. The least square fit to this data is shown in Fig. 4 (upper panel). Fitting to LMC data alone, or separately to EP or BD data yields similar results.

The existence of the $P-e$ relation could, in principle, stem solely from the detectability limit. The expression for the semi-amplitude, K, of the stellar radial velocity, induced by a companion orbiting a star with mass $M_\star$, can be written in the following form,

$$\mathcal{F}(K, M\sin i, P, e) = K - \left(\frac{2\pi G}{P}\right)^{1/3} \frac{M\sin i}{(M_\star + M)^{2/3}} \frac{1}{(1-e^2)^{1/2}} = 0 \qquad (6)$$

According to Marcy & Butler (1998), a confident detection requires that the semiamplitude be ~ 4 times the Doppler error, or $K_{\rm min} \approx 40\,{\rm m\,s^{-1}}$. Thus, a companion with a given projected mass $(M\sin i)_0$ can only be detected if its (P_0, e_0) is located above the curve $\mathcal{F}(K_{\min}, (M \sin i)_0, P, e) = 0$, on the $P-e$ diagram. Three examples of such curves are plotted on our $P-e$ diagram. We applied this criterion to all LMC and have found that the only object with a location on the $P-e$ diagram near its detectability limit is HD 210277 with $M \sin i = 1.36 M_{\rm J}$, $P = 437$ days, and $e = 0.45$. Thus, the existence of the $P-e$ relation is not an artifact of the detectability limit; the observed LMC (with the possible exception of HD 210277) could, in principle, have lower eccentricities and still be detectable. The general absence of LMCs with low eccentricities for periods in excess of a few tens of days is remarkable.

Fig. 4 (lower panel) is the often discussed $M \sin i - e$ diagram. The insert lists correlation coefficients for various subgroups of the data with $e > 0.001$. The entries in brackets give correlation coefficients for the data excluding additional objects in orbits suspected of being altered by stellar tides and thus having eccentricities lower than the nominal values. Excluded objects are υ Andromedae ($P = 4.621$ days), 51 Pegasi ($P = 4.2308$ days), and HD 283750 ($P = 1.79$ days). The statistical significance of $\rho_{\rm s} \neq 0$ is $s = 0.02(0.17)$ for EP, $s = 0.84(0.62)$ for BD, and $s = 0.05$ for stellar companions. To be detectable, the location, $((M \sin i)_0, e_0)$, of a companion with the period P_0 on the $M \sin i - e$ diagram must be above the curve $\mathcal{F}(K_{\min}, M \sin i, P_0, e) = 0$. Three examples of such curves are plotted on our $M \sin i - e$ diagram. Only HD 210277 is at the limit of detectability, and thus the $M \sin i - e$ diagram is not altered by the detectability limit.

The association between $M \sin i$ and e in the LMC and stellar companions populations is weak to non-existent. Most likely, these quantities are uncorrelated. Contrary to earlier claims (Mayor et al. 1998) the LMC population cannot be divided into two sub-populations on the basis of orbital eccentricity.

Overall, our calculations suggest that the LMC population is homogeneous with respect to statistics of orbital elements. EPs and BDs share a common PDF for orbital periods and eccentricities, they also share a common period-eccentricity correlation, as well as the same lack of significant mass-eccentricity correlation. Furthermore, the entire LMC population displays orbital elements statistics very similar to that of compatible stellar companions.

6. Discussion and conclusions

We have conducted a statistical analysis of LMC projected masses and orbital elements in an effort to assess whether the existing data provide unambiguous evidence for the presence of two populations of objects in the LMCs. Two of the four arguments presented to assert that there are two populations are beyond the scope of this paper, but the two central arguments involving the mass distribution of LMCs and a possible correlation between the mass and orbital eccentricities of the LMCs have been tested. Our findings are as follows:

(1) With respect to the projected mass distribution of LMCs, we have found indeed that there exist heterogeneous models that offer statistically significant better fits to the available data than a homogeneous model. In other words, there *is* an indication of a break in the projected mass distribution of LMCs. However, we also have found that there are at least two families of such heterogeneous models, set apart by the location of the power-law break and values of power-law indices that offer comparable fits to the data. One such family of models is compatible with the usual concept of EP and BD, but the other is not. Moreover, the scarcity of data makes it likely that the superiority of one or both types of heterogeneous models is due to a particular sample realization, and not necessarily indicative of the actual mass distribution. The best fit to a single power-law model has $p \approx 1$.

(2) Adjusting data for sample size and instrumental precision does not alter qualitatively the overall result. It does result in slightly steeper power-law in the case of a homegeneous model.

(3) All LMCs have the same orbital eccentricity and orbital period distribution functions. A homogeneous model is strongly suggested.

(4) The only clear correlation among LMC observables is between eccentricity and period. One cannot divide LMC into two sub-populations on the basis of orbital eccentricity as previously claimed.

(5) There is a striking populational similarity between LMCs and compatible stellar secondaries. The underlying eccentricity and period distribution functions, as well as correlations, for LMCs are indistinguishable from those constructed for stellar secondaries in the Heacox (1999) sample. Moreover, the distribution of projected masses of these secondaries is best approximated by a single power-law with an index of $p \approx 1$, a value about the same as that obtained while fitting a single power-law to the LMCs data. It cannot be overstressed that this does not require that LMCs and stellar secondaries share the same, *monotonically decreasing*, $\sim M^{-1}$, mass function, only that the pieces of the mass function in the domains of stellar secondaries and LMCs have the same functional form.

These results, taken all *together* put in question the prevalent assertion that the present data demonstrate existence of EP and BD as separate populations. In general, our findings are in agreement with those of Heacox (1999) who performed a similar analysis using a different statistical method. Use of empirical CDFs distinguish ours and Heacox's analysis from previous assessments which relied on histograms, a very subjective technique especially in the case of a small sample. However, it is also important to recall that our (as well as all previous) statistical analysis cannot be consider definitive because of possible bias of available LMC sample despite adjustments (see Sect. 2).

Apart from the possible existence of a break in the projected mass distribution of LMCs, all other evidence suggests a homogeneous population, possibly somehow related to the population of stellar companions. Populational similarity between LMCs and stellar secondaries is in striking contrast to marked dissimilarity between statistics of LMCs and those of secondary objects believed to have formed via accretion in a circumprimary disk. Such objects are the planets in our own planetary system, the regular satellites of the planets Jupiter, Saturn, and Uranus, and possibly the companions to the pulsar PSR 1257+12 (Wolszczan & Frail 1992).

Our findings invite questions about the origin of LMCs. The surprising populational similarity in orbital elements between EP and BD on the one hand, and the entire LMCs population and binaries on the other hand, implies a common cause or causes. The observed distributions can be functions of evolution (see discussion in Heacox 1998 and Heacox 1999), formation mechanism, or a combination of formation and evolution (Black 1997). Common formation mechanisms for all LMCs and some stellar secondaries is certainly a viable hypothesis, even if further observations strengthen the evidence for a break in the mass distribution, as long as the distributions of orbital elements noted here remains.

Note that the break in the mass distribution (if real) is, by itself, insufficient evidence for asserting a dual origin for LMCs. For example, it is thought that all planets in the solar system have a common accretionally based origin, and yet constructing a histogram of planetary masses would reveal a division into terrestrial and giant planets.

Given the present observations, a common origin hypothesis has no less merit than the prevalent hypothesis according to which EP formed in a process fundamentally different (i.e., as do planets) from BD and stellar companions. The term "common origin" is here taken to imply a similar, but not necessary identical, set of processes. We discuss one possible example later in this section.

In addition to being supported by the available orbital elements observations, the common origin hypothesis also has the virtue of simplicity. If we look at the LMC population from this perspective its allegedly "peculiar" properties suddenly look very ordinary. The location of Jupiter-size objects at very close distances from stars and moving in elliptical orbits are "natural" in a population related to stellar companions. Perhaps, the peculiarity of the objects popularly known as "extrasolar planets" is only due to misconception about their origin.

This simplicity is in contrast to what is required in the framework of the EP hypothesis. Orbital migration mechanisms have to be invoked to account for location of EP, but as such mechanisms would sweep a planet into the star, the addition of stopping processes is necessary. In addition many different mechanisms have been proposed to generate the observed eccentricities. (References to some of the work on the aforementioned mechanisms can be founded in Marcy et al. 1999). Although individually each of these mechanisms is theoretically viable, as a set, they have to be viewed as an adroitly chosen construct to support the standard planet hypothesis. Also, it is not clear how to account for giant planets in the Solar System in the framework of such a schemata! On the other hand, in the common origin hypothesis there is only a single outstanding problem of how to form Jupiter-mass analogs to stellar companions.

Interestingly, an idea, rooted in the planetary origin, but formally belonging to the category of the common origin hypothesis, has been proposed by Artymowicz et al. (1998). They discussed a possibility that all LMC began as planets. Their numerical calculations suggest that a growing protoplanet, while causing a gap in a disk, can continue to accumulate mass and grow to perhaps a brown dwarf-size object – a superplanet. To support their scenario, Artymowicz et al. (1998) pointed out that disk-planet interactions would naturally lead to superplanets having large eccentricities and regular planets having small eccentricities as in $M \sin i - e$ relation based on circa 97 data. However, as noted above, the current data do not support such division. Moreover, it is not clear how this scenario can account for the $P - e$ relation discussed here and in Black (1997). In addition, it is difficult to see how the LMC population formed that way can acquire statistical properties virtually identical to those of stellar companions unless we are willing to extend superplanets all the way to stellar masses.

The theoretical challenge is to come up with a feasible scenario for the common origin hypothesis. Here we offer some preliminary thought on one such scenario. Adams & Benz (1992) considered the possibility of forming binary companions by means of gravitational instabilities in circumstellar disks. Their scenario works as follows. At some early stage the disk mass is comparable to the stellar mass, which at that stage is much smaller than its final mass. Under such conditions, gravitational instabilities occur leading to the formation of a Jupiter-mass companion around a small star. Subsequent infall augment both the star and its companion to produce a typical binary system. According to Adams & Benz, this mechanism can, in principle, form binaries with separations anywhere in the range from the stellar radius to 100 AU. The character of such binaries depends on initial condition, timing, and a manner in which subsequent infall material is shared between the star and its companion. Perhaps, under most conditions, stellar binaries form, but under certain, less likely conditions or differing circumstances, LMCs form. At present, this scenario is only a suggestion that has to be considered more closely.

The possibility that the origin of some or all of the low-mass component of the LMCs is not a standard planet formation mechanism brings the issue of what the name "planet" signifies. We suggest that a definition of the term "planet" should center on how they are formed. Thus, if the further studies confirm that LMC form via a process fundamentally different from what is currently accepted, perhaps we should rethink calling them extrasolar planets.

In summary, in addition to presenting the results of our statistical analysis, the major goal of this paper is to raise the awareness of the fact that, although intellectually fascinating, the standard planetary hypothesis is not the only possible hypothesis for the origin of the LMCs. Based on statistics of LMCs, the common origin hypothesis is a viable alternative to the EP hypothesis. Because such a hypothesis was not presented before, its theoretical underpinning are not yet well developed, but this should change, especially if new observations continue to support our statistical findings.

Acknowledgements. We wish to thank W. D. Heacox for providing a preprint of his paper and binary data. We are also grateful to anonymous referees for helpful suggestions and comments on earlier draft of this paper. This research was conducted at the Lunar and Planetary Institute, which is operated by the Universities Space Research Association under contract No. NASW-4574 with the National Aeronautics and Space Administration. This is Lunar and Planetary Institute Contribution No. 1005.

References

Adams F.C., Benz W., 1992, In: McAlister H.A., Hartkopf W.I. (eds.) ASP Conf. Ser. 32, Complementary Approaches to Double and Multiple Star Research. ASP, San Francisco, p. 185

Artymowicz P., Lubow S.H., Kley W., 1998, In: Celnikier L.M., Trân Thanh Vân J. (eds.) Planetary Systems - the long view. Editions Frontieres

Black D.C., 1997, ApJ 490, L171

Boss A.P., 1998, ApJ 503, 923

Dempster A.P., Laird N.M., Rubin D.B., 1977, J. R. Statist. Soc. B 39, 1

Duquennoy A., Mayor M., 1991, A&A 248, 485

Gonzalez G., 1998, A&A 334, 221

Gonzalez G., Wallerstein G., Saar S.H., 1999, ApJ 511, L111

Heacox W.D., 1998, AJ 115, 325

Heacox W.D., 1999, submitted to ApJ

Mack C., 1967, Essentials of Statistics for Scientists and Technologists. Plenum Press, New York

Marcy G.W., Cochran W.D., Mayor M., 1999, to appear in: Mannings V., Boss A., Russell S. (eds.) Protostars and Planets IV

Marcy G.W., Butler R.P., 1998, ARA&A 36, 57

Mayor M., Queloz D., Udry S., Halbwachs J.-L., 1997, In: Cosmovici C.B., Bowyer S., Werthimer D. (eds.) IAU Colloq. 161, Bioastronomy 96, Editride Compositori, Bologna, p. 313

Mayor M., Queloz D., Udry S., 1998, In: Celnikier L.M., Trân Thanh Vân J. (eds.) Planetary Systems - the long view. Editions Frontieres

Mazeh T., Goldberg D., Latham D.W., 1998, ApJ 501, L199

Rees M.J., 1976, MNRAS 176, 483

Silk J., 1977, ApJ 214, 152

Wolszczan A., Frail D.A., 1992, Nat 355, 145