1. The Hi-GAL survey and goals of this work

Hi-GAL is a survey of the galactic plane in the range |l|<60° and |b|<1°, making use of PACS and SPIRE in parallel mode. Images of the continuum emission at 70, 160, 250, and 500 μm have been obtained, with angular resolution from 5" to 36" (Molinari et al. 2010).

The relationship between mass distributions of dense cores forming regions (also known as the mass function, MFF, or CMF) and the stellar IMF contains information regarding how observed samples of cores evolve into stars. The observational similarity between the CMF and the IMF has already been discussed in countless papers. The qualitative similarity between the CMF and the IMF offers support for the widely accepted idea that stars form from dense cores. Observed CMFs are typically characterized by one or more power-law distributions.

2. Bayesian statistical framework

The mass and mass-normalized mass distributions are described by the following probability density functions (PDF):

\[ P_{\text{mun}}(m) = C_{\text{mun}} m^{-\alpha} \]

Bayesian inference can be used for estimating the model parameters vector \( \mathbf{q} \) (i.e., \( C_{\text{mun}}, m_1, m_\text{inf}, \alpha \)) for power-law models. The normalization constants, \( C_{\text{mun}} \) and \( C_{\text{mun,in}} \), depend on \( m_0 \), the lower cutoff of the CMF, and for comparing these two competing models (\( m_{\text{inf}} \) and \( m_0 \)). Bayes’ theorem allows to calculate the ratio of the probabilities that the two competing models have generated the data \( D \), given any prior information \( I \), through the Bayes’ factor, \( B_{\text{p}}(D,I) \), which can be written as:

\[ B_{\text{p}}(D,I) = \frac{P(D|M_1,I)}{P(D|M_0,I)} = \frac{P(D|\mathbf{q})P(\mathbf{q}|I)}{P(D|\mathbf{q})P(\mathbf{q}|I)} \]

where \( P(D|M,I) \) is the likelihood of the data with parameters \( q \), and \( P(\mathbf{q}|I) \) is the prior probability, which represents the probability distribution of the unknown parameters \( q \) (Gregory, 2005).

Estimating the multi-dimensional integrals in Eq.(2) is impossible to be done analytically in most cases, and is otherwise computationally very intensive when done numerically.

3. Markov Chain Monte-Carlo

A mathematical tool able to efficiently evaluate the multi-dimensional integrals in Eq.(2) is the Markov chain Monte Carlo (MCMC) method.

The basic idea behind Monte Carlo integration is the following: we need to compute integrals (like in Eq.(2)) of the form \( \int f(x)dx \), where \( f(x) \) is a function of a random variable \( X \) which has probability density \( P(x) \). The MC integration thus replaces the integral with:

\[ \int f(x)P(x)dx = \frac{1}{N} \sum_{i=1}^{N} F(x_i) \]

where \( x_i \) are independent samples drawn from \( P(x) \). The MCMC method has the ability to generate a set of \( N \) independent samples \( x_1, x_2, \ldots, x_N \) whose distribution, after some finite number of steps (or “burn-in period”), is equal to \( P(x) \), thus allowing the integration.

Various methods exist to implement MCMC. The one used in this work is known as the Metropolis-Hastings method (Metropolis et al. 1953, Hastings 1970). A modified version of this algorithm was launched in parallel \( N \) to the original Metropolis-Hastings algorithm. In the Metropolis-Hastings method several versions of the MH algorithm are launched in parallel, each characterised by a different temperature parameter \( \beta \). MHM allows to visit regions of parameter space containing significant probability, not accessible to the basic algorithm.

5. Parameters Estimation with MHpt: POWERLAW case

Fig. 2. Distributions of the parameters \( m_1 \), \( m_\text{inf} \), and \( \alpha \) for the lognormal case in the l-60° field. We always employ a Jeffreys’ prior on parameter \( a \). (a) We adopt Jeffreys’ prior on \( m_\text{inf} \). (b) We employ a Gaussian prior on \( m_\text{inf} \). This example shows the importance of the inclusion of prior information. By incorporating our prior knowledge (see panel 4), we are able to yield a much cleaner constraint on \( m_\text{inf} \), which in panel (a) has a considerable tail of probability towards large values.

6. Parameters Estimation with MHpt: LOGNORMAL case

Fig. 3. Distributions of the parameters \( m_1 \) and \( m_\text{inf} \) for the lognormal case in the l-60° field. We employ a Jeffreys’ prior on \( \alpha \) and Gaussian priors on \( m_1 \) and \( m_\text{inf} \). It can be shown that the log global likelihood (see Eq.(2)) of a particular model \( M \) is given by (Gregory 2005):

\[ \ln P(D|M,I) = \frac{1}{N} \sum_{i=1}^{N} \ln P(D|M_i,I) \]

where

\[ \ln P(D|M_i,I) = \mu \ln P(D|M_i,I) \]

is the expected value of log-likelihood for a particular tempered MCMC chain characterized by \( \beta \). The set \( \{m_1\} \) represents the corresponding mass samples drawn from the MCMC chain with temperature parameter \( \beta \), while \( N \) is the number of samples in each set. Our current conclusion is that by using Eq. (4) and (5) we find that the Bayes factors, \( B_{\text{p}}(D,I) \), favoring the power-law. However, the prior effects and other computational details of the MHpt method are still being analyzed.

References

Hastings, W.K., 1970, Biometrika, 57, 97
Molinari, S. et al., 2010, A&A 518, L100
1 INAF-OAA, 2 IPR-RP, 3 Univ. of Arizona, 4 INAF-IFSI, 5 CESR, 6 OAMP, 7 Univ. of Kent, 8 ESO, 9 Univ. of Hertfordshire

L. Olmi1,2, A. Anglés-Alcázar3, D. Elia4, S. Molinari4, L. Montier5, M. Pestalozzi6, S. Pazzuto4, D. Polychroni4, R. Ristorcelli5, J. Rodon6, M.D. Smith7, L. Testi8, and M. Thompson9

1 INAF-OAA, 2 IPR-RP, 3 Univ. of Arizona, 4 INAF-IFSI, 5 CESR, 6 OAMP, 7 Univ. of Kent, 8 ESO, 9 Univ. of Hertfordshire

2. Bayesian statistical framework

The mass and mass-normalized mass distributions are described by the following probability density functions (PDF):

\[ P_{\text{mun}}(m) = C_{\text{mun}} m^{-\alpha} \]

Bayesian inference can be used for estimating the model parameters vector \( \mathbf{q} \) (i.e., \( C_{\text{mun}}, m_1, m_\text{inf}, \alpha \)) for power-law models. The normalization constants, \( C_{\text{mun}} \) and \( C_{\text{mun,in}} \), depend on \( m_0 \), the lower cutoff of the CMF, and for comparing these two competing models (\( m_{\text{inf}} \) and \( m_0 \)). Bayes’ theorem allows to calculate the ratio of the probabilities that the two competing models have generated the data \( D \), given any prior information \( I \), through the Bayes’ factor, \( B_{\text{p}}(D,I) \), which can be written as:

\[ B_{\text{p}}(D,I) = \frac{P(D|M_1,I)}{P(D|M_0,I)} = \frac{P(D|\mathbf{q})P(\mathbf{q}|I)}{P(D|\mathbf{q})P(\mathbf{q}|I)} \]

where \( P(D|M,I) \) is the likelihood of the data with parameters \( q \), and \( P(\mathbf{q}|I) \) is the prior probability, which represents the probability distribution of the unknown parameters \( q \) (Gregory, 2005).

Estimating the multi-dimensional integrals in Eq.(2) is impossible to be done analytically in most cases, and is otherwise computationally very intensive when done numerically.

4. Prior Information

A possible procedure to obtain prior information to be used with Bayesian inference techniques, is to construct a histogram representing the mass frequency distribution, and then extract the parameters \( q \) by performing a least-squares linear regression on a log-log histogram. Unfortunately, this method can generate significant systematic errors.

We therefore use a procedure (Clauset et al. 2009) that implements the maximum likelihood estimator (MLE) for fitting the power-law to data distribution, along with a goodness-of-fit based approach to estimating \( m_{\text{inf}} \). This procedure (PLFIT) allows us to set constraints on the values of \( m_{\text{inf}} \) and \( \alpha \). For the lognormal distribution we still perform a least-squares regression on the log-log histogram.

Fig. 1: The left panel shows a standard lognormal fit to a log-log histogram reproducing the CMF for the Hi-GAL 60°-90° field. On the right panel the same CMF is shown (binning is different) with overlaid the line with the slope as given by the PLFIT method, which is independent from binning.