Astronomy & Astrophysics

Deriving structural parameters of semi-resolved star clusters

FitClust: a program for crowded fields

D. Narbutis[1], D. Semionov[1], R. Stonkutė[1], P. de Meulenaer[1,2], T. Mineikis[1,2], A. Bridžius[1,2], and V. Vansevičius[1,2]

[1] Center for Physical Sciences and Technology, Savanorių 231, 02300 Vilnius, Lithuania
e-mail: donatas.narbutis@ftmc.lt
[2] Vilnius University Observatory, Čiurlionio 29, 03100 Vilnius, Lithuania

ABSTRACT

Context. An automatic tool to derive structural parameters of semi-resolved star clusters located in crowded stellar fields in nearby galaxies is needed for homogeneous processing of archival frames.
Aims. We have developed a program that automatically derives the structural parameters of star clusters and estimates errors by accounting for individual stars and variable sky background.
Methods. Models of observed frames consist of the cluster's surface brightness distribution, convolved with a point spread function; the stars, represented by the same point spread function; and a smoothly variable sky background. The cluster's model is fitted within a large radius by using the Levenberg-Marquardt and Markov chain Monte Carlo algorithms to derive structural parameters, the flux of the cluster, and individual fluxes of all well-resolved stars.
Results. FitClust, a program to derive structural parameters of semi-resolved clusters in crowded stellar fields, was developed and is available for free use. The program was tested on simulated cluster frames, and was used to measure clusters of the M31 galaxy in Subaru Suprime-Cam frames.
Conclusions. Accounting for bright resolved stars and variable sky background significantly improves the accuracy of derived structural parameters of star clusters. However, their uncertainty remains dominated by the stochastic noise of unresolved stars.

Key words. galaxies: star clusters: general – methods: data analysis – techniques: image processing

1. Introduction

An automatic tool to derive structural parameters of star clusters is essential so that numerous archival frames from stellar population surveys in nearby galaxies can be processed homogeneously. Observed surface brightness distribution in frames is commonly treated as a sum of the sky background and the cluster's profile, e.g., in *Hubble* Space Telescope (HST) studies of clusters in galaxies: M 33 (San Roman et al. 2012), M 31 (Tanvir et al. 2012), and NGC 7252 (Bastian et al. 2013), or ground based observations of M 31 clusters (Barmby et al. 2007). However, the aforementioned studies are focused mostly on bright clusters. For fainter clusters, located in crowded fields of galaxy disks, the brightest individual stars have to be taken into account properly in order to avoid bias in structural parameters, derived by fitting cluster models to the surface brightness distribution.

We propose a method to solve an inverse problem aimed at finding best-fit for 1) a cluster; 2) an unknown number of bright stars (cluster members or foreground objects); and 3) a variable sky background.

Structural parameters of clusters are degenerate and are due to problems of a sky background determination in crowded areas (Werchan & Zaritsky 2011). It is virtually impossible to determine a true sky background in a crowded field frame because of unresolved sources of variable spatial number density. For some clusters in the M31 galaxy (Narbutis et al. 2008), which reside on dust lanes, severe background variability was observed. The strongest effect due to the sky background determination

inaccuracy is on the parameters describing luminosity distribution in the outskirts of clusters, e.g., tidal radius.

In a study of a semi-resolved star cluster sample compiled by Kodaira et al. (2004) from a Subaru Suprime-Cam survey of the M31 galaxy, a grid of model surface brightness profiles convolved with the point spread function (PSF) was constructed and used to derive parameters of the observed clusters (Šablevičiūtė et al. 2007). However, this method did not take into account individual stars and variability of the sky background.

Larsen (1999) developed a program ISHAPE to derive structural parameters of unresolved clusters based on masking out deviating pixels. It was applied to the study of the M31 clusters (Šablevičiūtė et al. 2006). However, for extended objects significantly larger than the size of the PSF, stars could not be masked out successfully.

Frames of crowded fields are usually processed using DAOPHOT (Stetson 1987) or similar programs to make stellar photometry catalogs. However, they have a resolution limit, and split unresolved cores of clusters into several artificial point sources. Hence, they are suitable for construction of color−magnitude diagrams of the brightest stars only in the outer parts of semi-resolved clusters, e.g., an extended object studied by Stonkutė et al. (2008) or clusters measured in HST frames by Larsen et al. (2011).

Therefore, a robust tool suitable for deriving cluster parameters in crowded stellar fields is needed.

In this paper we present a program, FitClust, designed to derive structural parameters of clusters with overlapping bright

stars located on a variable sky background. FitClust is implemented in Python and uses DAOPHOT within PyRAF. The code can be obtained from the FitClust website[1].

In Sect. 2 we describe a method implemented in the program; in Sect. 3 we analyze its performance with artificial clusters and apply it to the study of clusters in the M31 galaxy.

2. Deriving structural parameters

In order to derive structural cluster parameters by employing FitClust, the following input data are required:

- an observed frame containing a cluster;
- a normalized PSF (i.e., the sum of pixel values is set to unity) derived for the frame;
- a bad pixel mask;
- a model of a cluster as an analytic 2D brightness distribution (coordinates, shape parameters, and total flux);
- coordinates and fluxes of the resolved stars identified and fitted with the PSF;
- a sky background: a fixed value (for the flat sky) or an approximation with a 2D spline (for the variable sky);
- a set of algorithm control parameters.

The basic steps of the iterative cluster and star fitting algorithm implemented in FitClust are displayed in Fig. 1. The main loop of model fitting consists of 1) measurement and subtraction of sky background; 2) fit of the cluster and stars together; 3) subtraction of the best-fit cluster to produce a residual frame; 4) derivation of more accurate coordinates of stars on the residual frame; 5) subtraction of the cluster and stars to measure sky background for the next iteration; and 6) sampling of parameter space with Markov chain Monte Carlo (MCMC) algorithm.

The main loop runs for a given number of iterations and produces the best-fit cluster model parameters, which can optionally be fitted using MCMC algorithm to derive their uncertainties. The parameter values of model components are derived by minimizing an array of residuals within specified model fitting area, weighting each pixel by its uncertainty assuming photon noise.

2.1. Input frames

FitClust takes a small subframe from the observed frame centered on a cluster and uses it as an input. If the frame contains CCD defects or saturated stars, they are excluded by using a bad pixel mask, which can be created automatically (e.g., with the program StarL by Narbutis et al. 2009) and edited manually.

Usually, there are several clusters to be measured in the observed frame. For convenience, FitClust can automatically construct the empirical PSF for the entire frame using the DAOPHOT program within the IRAF system (Tody 1993). If the PSF is strongly variable, which is the case for wide-field imaging, the PSF for individual clusters (subframes) can be constructed based on isolated stars located in the cluster's vicinity.

2.2. Sky background determination

A sky background value is determined for every iteration (Fig. 1) of cluster model fitting and subtracted from the input frame. For the first iteration, the original input frame is used to construct a sky value histogram, whose lower pixel value part is fitted with a

Gaussian function to derive the initial sky level value. For subsequent iterations, a residual frame with the cluster and stars subtracted is used to measure a sky background either by 1) fitting a sky value with a Gaussian function; or 2) fitting a 2D spline function.

2.3. Parameters

The following parameters are used to control program execution: 1) a type of cluster model surface brightness distribution; 2) approximate initial coordinates (identified by eye, by computing mass center of flux distribution using moments, or by fitting a 2D Gaussian function) of a cluster center, x_{in}, y_{in}; 3) a radius of a circle defining the fitting area of the model, r_{fit}; 4) a number of fitting and star detection iterations, n_{fit}; and 5) the limits for fitting parameters.

A projection of an unresolved cluster on the frame is usually described by tidal-cut-off King (King 1962) or power-law EFF (Elson et al. 1987) profiles, which are used to simulate a smooth surface brightness distribution. A frame of a cluster model is constructed using values of an analytic function at the center of each pixel and normalizing flux to unity. For the profiles of the EFF and King models, a truncation radius is used for normalization.

The King surface brightness profile is defined by a central surface brightness, μ_0; a core radius, r_c; and a tidal radius, r_t:

$$\mu(r) = \mu_0 \left[\left(1 + \frac{r^2}{r_c^2} \right)^{-1/2} - \left(1 + \frac{r_t^2}{r_c^2} \right)^{-1/2} \right]^2. \qquad (1)$$

The EFF profile is defined by μ_0; a scale radius, r_e; and a power-law index, n:

$$\mu(r) = \mu_0 \left(1 + \frac{r^2}{r_e^2} \right)^{-n}. \qquad (2)$$

A cluster model is described by center coordinates, x_c, y_c; two constants defining a shape of the profile, p_1, p_2; and a total flux, f_c, which is linked to the central surface brightness, μ_0. Analytic brightness distribution is convolved using FFTW library[2] with the PSF to mimic observation effects. Stars are represented by the PSF shifting it to the desired positions, x_i, y_i, using spline interpolation and scaling by fluxes, f_i.

Therefore, the following cluster model parameters are provided:

- x_c, y_c – center coordinates of a cluster;
- p_1, p_2 – for the King model: the core radius, r_c, and the ratio of tidal to core radii, r_t/r_c; for the EFF model: the scale radius, r_e, and the power-law index, n;
- f_c – the total flux of a cluster;
- x_i, y_i, f_i – center coordinates and the flux of an ith star;
- μ_{sky} – the sky background: constant or approximated by a 2D spline function.

2.4. Model fitting

The loop of model fitting shown in Fig. 1 can be summarized as a following sequence of steps repeated for a given number of iterations, n_{fit}:

- measure sky background, μ_{sky}, on 1) an input frame in the first iteration; 2) a residual frame (with a cluster and stars

---

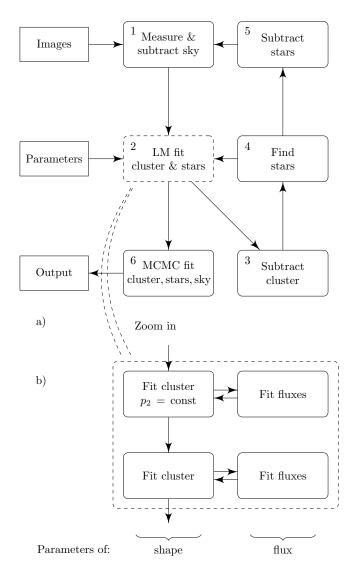[1] FitClust: www.astro.ff.vu.lt/software/fitclust

[2] http://www.fftw.org

Fig. 1. Schematic representation of the iterative cluster and star fitting algorithm with FitClust. Panel **a)** shows the basic steps and data blocks, while panel **b)** elaborates on the innermost loop of cluster parameter fitting, which are separated into two groups.

subtracted) in subsequent iterations; subtract sky background from the input frame;

– derive cluster parameters x_c, y_c, p_1, p_2 by 1) fitting only a cluster in the first iteration; 2) fitting cluster and stars in the subsequent iterations while keeping their coordinates, x_i, y_i, fixed (derived during previous iteration); derive flux of a cluster, f_c, and of stars, f_i, by performing algorithm steps defined in Fig. 1b;

– make a residual frame by subtracting fitted cluster model from the input frame;

– use DAOPHOT on the residual frame to find resolved stars and measure their coordinates, x_i, y_i; pass list of star coordinates to the next iteration of model fitting;

– subtract stars from the residual frame and pass it to the next iteration of sky background determination.

To derive parameters of a cluster as shown in Fig. 1b, we use the Levenberg-Marquardt (LM) algorithm[3] and constrain parameter limits. We have found that fitting cluster coordinates, x_c, y_c, and

---

shape parameters, p_1, p_2, simultaneously does not perform correctly, because solutions get stuck in local minima. Therefore, we run the first pass of LM algorithm and fit only x_c, y_c, p_1, while $p_2 = $ const. (King $r_t/r_c = 10$ and EFF $n = 1.5$). The coordinates and both shape parameters are fitted in the second pass of LM algorithm.

The model frame is constructed from a cluster model and a reasonable number of stars ($N \lesssim 20$, because time of fitting is proportional to N^2), all stored as individual 2D arrays with volumes normalized to unity. For each pixel of a residual frame we have

$$\mu_{res} = \mu_{in} - \mu_{sky} - \mu_{mod}, \tag{3}$$

$$\mu_{mod} = f_c \cdot \mu_c - \sum_{i=1}^{N} f_i \cdot \mu_{PSF}. \tag{4}$$

Here μ_{res} is a pixel value of the residual frame; μ_{in} of the input frame; μ_{sky} of the sky background; μ_{mod} of the model frame, which contains a cluster flux, f_c, multiplied by cluster model, μ_c; and a sum of star fluxes, f_i, multiplied by the PSF model, μ_{PSF}. For the LM fitting the fluxes of cluster and stars are constrained to be positive. Ideally, one would like to obtain residuals comparable to the measured sky background noise. Practically, stochastic noise of unresolved stars and variable crowding conditions within the cluster make residuals larger and dependent on the cluster's age, mass, and size.

2.5. Iterative detection of stars

Approximate coordinates of a cluster in the frame, x_{in}, y_{in}, are used to initialize the fitting procedure. Initially, we fit and subtract only a model of a cluster and pass the residual frame to DAOPHOT.

To find stars on the residual frame a `daofind` routine is used with a detection threshold of 4 standard deviations above the sky background level. Stars are measured with `allstar` and then subtracted. A second pass of `daofind` is performed to find additional stars in the residual frame. Lists of stars are concatenated and a second pass of `allstar` produces a list of star coordinates, x_i, y_i. Stars located within a fitting radius, r_{fit}, are selected and provided to the next iteration of the model fitting loop. Figure 2 shows a cluster frame decomposed into a cluster model, stars, and a residual frame.

In each iteration stars are identified anew and their coordinates are used only for subsequent iteration. Therefore, the number of stars, N, and their coordinates, x_i, y_i, vary and converge to the best-fit values. Cluster parameters and sky background also converge to the best-fit values, while the standard deviation of residuals, σ_{fit}, within r_{fit} decreases. The model fitting ends after a preselected number, n_{fit}, of iterations – typically n_{fit} is approximately 6.

2.6. Parameter fitting with MCMC

We have found that iterative model fitting performs well when detecting the brightest stars and deriving the parameters of old and massive clusters with smooth profiles. When young clusters of low mass are analyzed, fitting can get stuck in local minima because of stochastic variations of flux in the cluster's profile and stellar crowding.

To avoid local minima and to estimate parameter uncertainties, we use `emcee`, a Python implementation (Foreman-Mackey et al. 2013) of the affine-invariant MCMC ensemble sampler proposed by Goodman & Weare (2010). It starts an ensemble of
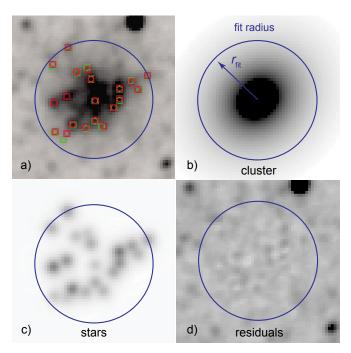
Fig. 2. Frame of one of the largest clusters from the M31 sample (Vansevičius et al. 2009) shown in panel **a)**; it is decomposed into cluster model **b)**, stars **c)**, and residual frame **d)**. In panel **a)** coordinates of stars detected in the first iteration are shown in green, the final iteration in red.

random walkers in the model parameter space from a position of the best-fit values of cluster, star, and sky background parameters obtained by the LM algorithm.

As each walker makes steps through the parameter space, a model of a cluster with a fixed number of stars is constructed at each step. We note that all cluster parameters, including star positions and fluxes, are allowed to vary when making MCMC steps. At each step a likelihood is computed for the input frame to be generated by the model with given parameter values, assuming that uncertainty of each pixel is Gaussian with a standard deviation equal to the square root of the input pixel's value.

After the burn-in phase, which takes ∼3000 steps when fitting a cluster without stars, `emcee` converges to the most likely parameter values and subsequent walker steps are confined within the limits of parameter uncertainties. The second sample of ∼1000 walker positions is analyzed. The median of the sample and its standard deviation are reported as the most likely value of parameter and its uncertainty.

For a cluster with one star included, ∼6000 steps are required for burn-in. The time needed to create a model image steeply increases with number of included stars. Therefore, practical considerations control the limit on the number of the brightest stars to include.

2.7. Output

The final step of the model fitting is numerical integration to derive a cluster's half-light radius, r_h. For this purpose a cluster model is produced with a frame resolution that is ten times higher. On this frame a curve of growth is constructed by integrating flux through circular apertures up to a truncation radius, r_{max}, where the surface brightness, μ, of a cluster model reaches 1% of its central value, μ_0.
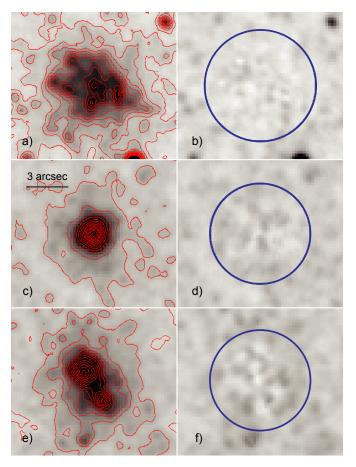
Fig. 3. *Left*: examples of M31 star clusters observed with Suprime-Cam taken from the study by Vansevičius et al. (2009), over-plotted with iso-flux lines, *from top to bottom*: semi-resolved, compact, and double. Image scale is indicated in panel **c)**. *Right*: residuals after subtracting the cluster model and stars; circles indicate fitting area.

In addition to the frames shown in Fig. 2, FitClust also outputs a record file containing parameters of a cluster model and stars, as well as σ_{fit}, which can be used to analyze the quality of fitting.

3. Tests and results

We have used FitClust to derive structural parameters of clusters in crowded fields of the M31 disk taken from the cluster sample presented by Narbutis et al. (2008). Examples of some characteristic objects are shown in Fig. 3. Parameters for a double cluster (see Fig. 3e) with two peaks in luminosity distribution were derived by fitting two cluster models simultaneously. In all three cases the subtracted cluster and stars produce a clean residual frame with a standard deviation, σ_{fit}, equal to that of sky background. Initially we attempted to fit elliptical models, but because of dominating stochastic effects in low mass clusters, analysis was limited to circular ones.

We simulated frames of artificial clusters and analyzed them with FitClust in order to evaluate its performance when detecting well-resolved stars automatically, determining sky background, and deriving cluster structural parameters. Artificial cluster frames resemble the quality of those obtained with Suprime-Cam in M31 studied by Vansevičius et al. (2009): 1) frame scale 0.2 arcsec/pixel; 2) full width at half maximum (FWHM) of the PSF 3 pixels.
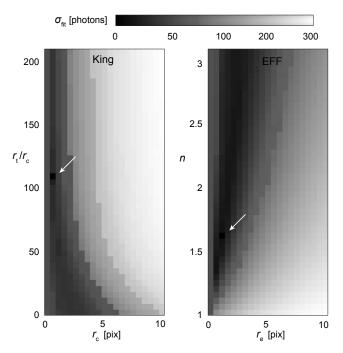
Fig. 4. Maps of fitting standard deviations, σ_{fit}, for artificial clusters based on smooth King (*left*) and EFF (*right*) models. Position of the cluster under consideration in the parameter space is indicated by the arrow ($r_c = 1$ pixel, $r_t/r_c = 110$, and $r_e = 1.5$ pixel, $n = 1.7$) and coincides with minimal value of standard deviation.

In the model fitting routine we set the following limits for cluster structural parameters: $r_c = r_e = [0.01, 15]$ pixels (both models), $r_t/r_c = [2, 500]$ (King), and $n = [1.01, 5]$ (EFF).

Just to illustrate that even in the simplest cases structural parameters are significantly degenerate, we simulated smooth artificial clusters using King and EFF models and convolved them with a PSF. Neither photon noise nor overlapping stars were included in this test. Fitting was performed at the nodes of the grid (p_2 vs. p_1) with both parameters fixed. The derived maps of fitting standard deviations, σ_{fit}, are shown in Fig. 4. A shallow band of σ_{fit} minimum is seen over a wide parameter range, where parameters are degenerate. These maps become even more complicated (have few local minima) when several stars are introduced into fitting.

3.1. Bright star test

We tested the algorithm to derive parameters of smooth model cluster in the presence of three resolved stars and uniform sky background, i.e., without faint unresolved objects.

The smooth clusters assuming the King and EFF models were simulated at the nodes of parameter grids displayed as black dots in Figs. 5 and 6, respectively. Nodes of the parameter grids were chosen to cover the parameter space of M31 clusters with $V < 20$ mag, taken from the dataset produced for the study by Vansevičius et al. (2009) and are shown with open circles in Figs. 5c and 6c. Flux of a model cluster was set to 10^6 photons, which corresponds to $V \sim 17.5$ mag for the object in the M31 cluster sample. The cluster model was convolved with a PSF and placed at the center of 101×101 pixel frame. At each node of the grid series of 100 frames were simulated.

Each cluster was superposed with three stars randomly placed within a circle of a 5 pixel radius centered on the cluster to make star detection more difficult. Figure 5a shows cases when

each interfering star has a flux of 10^5 photons. Since the total flux of three interfering stars makes up to $\sim 30\%$ of the cluster's flux, they alter the cluster luminosity distribution significantly.

Uniform sky background was modeled assuming a flux of 10^3 photons per pixel. The Gaussian photon noise was added to each pixel of the frame, assuming its standard deviation to be equal to the square root of the number of photons in a pixel. The cluster's fitting radius was set to 31 pixels; ten FitClust star detection and model fitting iterations were performed.

As can be seen in Fig. 5a, clusters with a small core radius tend to have larger measured r_c, and clusters with large r_c demonstrate an opposite effect. The tidal radius, r_t, decreases for all nodes, except for those with the largest r_c. Therefore, stars projected on the cluster introduce a significant bias into the derived cluster structural parameters. We note, however, that a similar effect on parameters, as shown in Fig. 5a, is also seen when even a single star is projected on the cluster. The three stars were chosen to demonstrate FitClust's ability to identify and measure stars projected in various configurations on the cluster.

When automatic detection of stars in FitClust is turned on, the derived structural parameters are in a good agreement with the input parameters of simulated clusters, see Fig. 5c larger ellipses around grid points. The scatter of the parameters shown by 1σ ellipses is smaller than in the cases when stars are not fitted (Fig. 5a). However, it is non-negligible because of photon noise and limited accuracy of star subtraction procedure. In most cases the parameters of stars are derived correctly. The flux of clusters is derived with an average accuracy of $\sim 5\%$.

The results of King model fitting in cases of three stars with flux of 4×10^4 photons each are shown in Fig. 5b. The total flux of stars makes $\sim 12\%$ of the cluster's flux and their disturbance of measured structural parameters is lower. However, directions of systematic shifts remain similar to those seen in Fig. 5a. Figure 5(c, smaller ellipses) shows derived parameters when stars are fitted together with the cluster. The ellipses are smaller than in Fig. 5b, because the residuals remaining after subtraction of fainter stars are less pronounced. The flux of clusters in these cases is derived with an average accuracy of $\sim 3\%$.

The results of cluster parameters derived from frames simulated with the smooth EFF model are shown in Figs. 6a and b for cases of interfering stars with flux of 10^5 and 4×10^4 photons, respectively. As in the case of King model, when stars are not fitted, scale radius, r_e, and power-law index, n, are significantly biased. However, automatic detection and fit of stars in FitClust switched on again gives correct structural cluster parameters (Fig. 6c), and it takes six iterations on average to obtain a consistent solution for stars and the cluster.

3.2. Sky background test

To test performance of FitClust in realistic sky background conditions, we selected two regions in the M31 Suprime-Cam CCD mosaic (*V*-band) taken from Narbutis et al. (2008). The first region represents a crowded field with ~ 20 stars within fitting radius. The second is located on the edge of the prominent dust lane, which introduces a sky background gradient across the 101×101 pixel frame used to derive structural cluster parameters.

Smooth cluster model frames were generated according to the prescription given in Sect. 3.1, except that they were convolved with the empirical PSF of Suprime-Cam. At each node of the parameter grids 100 cluster frames were generated and overlaid on the subframes of two selected M31 regions by randomly shifting coordinates within a box of 50 pixels. Model
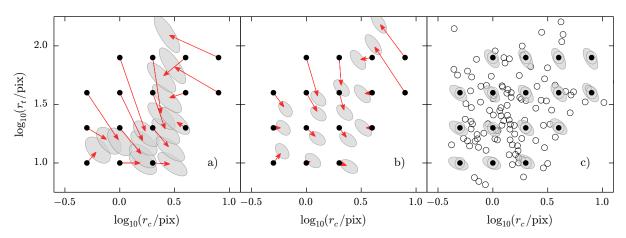
Fig. 5. Bright star test. Results of the King model fitting on frames of smooth artificial clusters, superposed by three identical stars. The total flux of the cluster is 10^6 photons. The cases of each star flux of 10^5 and 4×10^4 photons are shown in panels **a)** and **b)**, respectively. Panels **a)**, **b)** display cases when stars are not included in the model fit, and **c)** when they are automatically identified and fitted (smaller ellipses are for fainter stars originally in panel **b)**). Tidal radius, r_t, and core radius, r_c, are displayed. Dots mark nodes and indicate initial parameters; for each node series of 100 frames are simulated. Ellipses approximate 1σ scatter of the derived parameters and arrows show their bias. Open circles (panel **c)**) show M31 clusters from Vansevičius et al. (2009) with $V < 20$ mag.
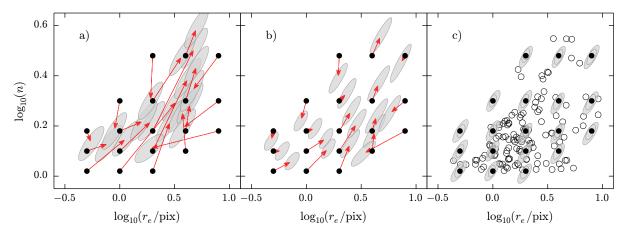

Fig. 6. Same as Fig. 5, but for smooth clusters simulated with the EFF model. The power-law index, n, and the scale radius, r_e are displayed.

fitting radius was set to 25 pixels and seven star detection and model fitting iterations were performed.

To estimate the accuracy of the derived cluster parameters versus cluster brightness, we have repeated the same test with artificial clusters having fluxes of 10^6, 4×10^5, and 1.6×10^5 photons, which correspond to approximate V-band magnitudes of 17.5, 18.5, and 19.5 mag (approximate mass range from 30 000 $M_\odot$ to 3000 $M_\odot$ for the clusters of 100 Myr age) in Suprime-Cam data.

The results of the King and EFF model fitting are shown in Figs. 7a and b, respectively. Three concentric 1σ ellipses, representing the derived parameter distributions around each node, correspond to the artificial clusters of V = 17.5, 18.5, and 19.5 mag, when real M31 stars and variable sky background are accounted for in model fitting. We note that in both cases of the real M31 environment (crowded and with sky background gradient) the distributions of derived parameters of each node look similar, therefore, combined results of both cases are shown in Fig. 7.

Arrows show systematic biases of the derived parameter distributions when stars and variable sky background are not taken into account in model fitting for clusters of V = 17.5 (red/shorter) and 18.5 mag (blue/longer). Significant shifts larger

than 2σ of the distributions are observed; however, they are even more severe in the case of V = 19.5 mag.

Comparing derived parameters of the models presented in Fig. 7 to the corresponding models in Figs. 5 and 6, we note that ellipses of the derived parameter distributions (in Fig. 7) are aligned more parallel to the vertical axes, indicating significant effect of the real sky background.

Correctly derived parameters of clusters (Fig. 7) suggest that FitClust can deal with ground-based images of crowded fields with variable sky background successfully for clusters with smooth profiles.

3.3. Stochastic cluster test

Previous tests used smooth cluster models, which are a good approximation of massive old clusters. Younger objects of lower mass are affected by stochastic luminosity fluctuations of stars and their random spatial distribution. To test performance of FitClust on them, we followed a method described by Larsen et al. (2011) to simulate clusters star-by-star and used the SimClust program Deveikis et al. (2008) to build V-band images of models located in M31 observed with Suprime-Cam.

We investigated two cluster model cases: 1) 100 Myr age with mass 10^4 $M_\odot$; and 2) 10 Gyr age with mass 10^5 $M_\odot$. They
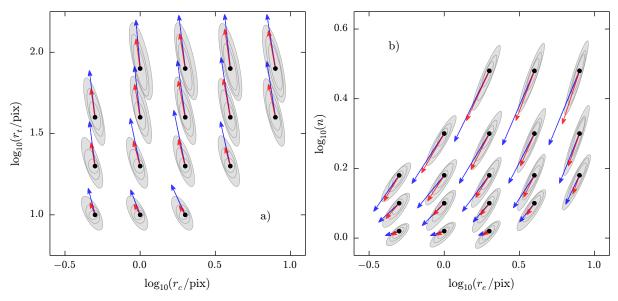
Fig. 7. Sky background test. Results of the King (panel **a**); tidal radius, r_t, vs. core radius, r_c) and EFF (**b**); power-law index, n, vs. scale radius, r_e) model fitting for artificial clusters placed in the real M31 Suprime-Cam CCD mosaic from Narbutis et al. (2008). Black dots indicate initial parameters. For each node series of 100 frames were simulated. Three concentric 1σ ellipses, representing the derived parameter distributions around each node, correspond to the artificial clusters of $V = 17.5$, 18.5, and 19.5 mag. Arrows show systematic biases of derived parameter distributions, when stars and variable sky background are not taken into account in model fitting for clusters of $V = 17.5$ (red/shorter) and 18.5 mag (blue/longer).

represent most of the young disk clusters and old globular clusters from the study by Vansevičius et al. (2009), and have typical fluxes of 5×10^5 and 1.3×10^5 photons in simulated images, respectively.

Stars were distributed according to the King model considering six combinations of core and tidal radius nodes with $r_c = 0.8, 1.5, 3.0$ pixels, and $r_t = 15, 40$ pixels (see Fig. 8). One hundred cluster images were generated per node. To estimate the influence of a bright star superposed on a cluster, two additional image sets were generated by placing a star (1 mag or 2 mag fainter than the typical total flux of a cluster) at a distance equal to the FWHM of a cluster from its center. Uniform sky background was modeled assuming a flux of 10^3 photons per pixel. The Gaussian photon noise was added to each pixel of the frame, assuming its standard deviation to be equal to the square root of the number of photons in a pixel.

Since the purpose of this test is to demonstrate influence of stochastic effects on derived structural parameters, we turned off automatic detection of stars and performed model fitting using MCMC sampling, starting from the given parameters of the simulated cluster and a bright superposed star, i.e., assuming that the position of the included star is known.

We note that the uncertainties of parameters reported by the MCMC for a single cluster are smaller than the size of dots indicating positions of grid nodes in Fig. 8. Therefore, they are significantly underestimated and do not show true uncertainties of parameters for clusters of given age and mass.

Ellipses in Fig. 8a approximate 1σ derived parameter scatter for 100 models per node of 100 Myr age, while Fig. 8f of 10 Gyr age clusters. The derived parameter values are centered around input value. However, uncertainties due to stochastic effects, especially of tidal radius, r_t, are two times larger for 100 Myr clusters than for 10 Gyr ones and are significant.

When a bright star was included in the image, a cluster was fitted in two ways: using only the cluster model (Figs. 8b, d, g, i), and including the star in the fit (Figs. 8c, e, h, j). If a background star, which is 1 mag fainter than the cluster, is superposed onto

100 Myr object (Fig. 8b), significant parameter bias (shown with arrows) is observed. A similar, but weaker effect is seen when the superposed star is 2 mag fainter than the cluster (Fig. 8d).

The same effect is observed for 10 Gyr clusters, although the largest objects in Fig. 8g show extreme bias because the cluster's center migrates to the position of the superposed star and core radius, r_c, decreases drastically, while the tidal radius increases to compensate for asymmetric luminosity distribution around it. Once a star is included in the model fitting, parameters of the cluster are recovered reliably (Figs. 8c, e, h, j); their scatter and values are close to the ones derived for cluster images without included star.

We have found that although the initial position of the input star is known, it is slightly adjusted (up to ~0.2 pixel and ~0.5 pixel for the 1 mag and 2 mag star cases, respectively) during the MCMC burn-in phase due to the underlying stochastic fluctuations on a 100 Myr cluster. We have made tests by allowing the initial position of the star to be offset and letting for the burn-in phase to adjust it. We have found that it has to be known with an accuracy of ~1 pixel for the MCMC fitting to converge to the input parameter values. Otherwise, when there are neighboring stars of comparable luminosity, i.e., in young clusters, a star migrates from its true position. Therefore, only the brightest stars can be considered in the MCMC fitting.

3.4. Star subtraction test

We performed a star subtraction test to quantify the influence of the cluster's brightest stars on the accuracy of its derived structural parameters. A set of cluster model images of 100 Myr, described in Sect. 3.3, was used; 1σ distributions of their recovered parameters are repeated in all panels of Fig. 9 as open ellipses centered on input parameter values.

We have modified the original images by subtracting the three brightest stars located anywhere in the cluster (i.e., in the core or in the outskirts) and derived their parameters, which
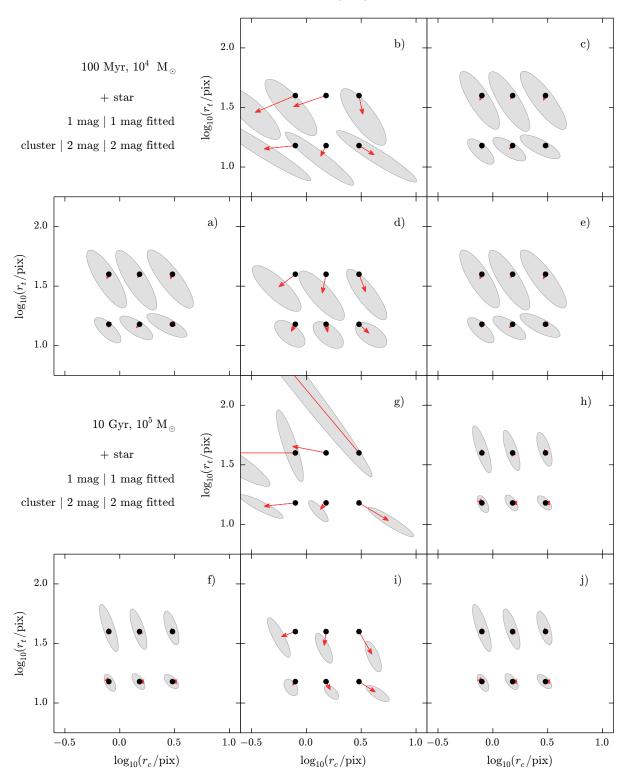
Fig. 8. Stochastic cluster test. Panels show King model clusters of age and mass: **a)**–**e)** 100 Myr, $10^4 M_\odot$, **f)**–**j)** 10 Gyr, $10^5 M_\odot$. In each panel dots indicate input r_c and r_t values of six nodes, ellipses approximate 1σ scatter of the derived parameters of 100 models per node, and arrows show their bias. Panels **a)** and **f)** show original models. Panels **b)** and **g)** are cases where the cluster is superposed with a star 1 mag fainter than the cluster, **d)** and **i)** – 2 mag fainter. Star was placed at a distance equal to FWHM of cluster from its center. Panels **c)**, **e)**, **h)**, and **j)** have the same star superposed (see corresponding row), but here the star is included in the model fit.

are displayed in Fig. 9a. This was repeated by subtracting the ten brightest stars; Fig. 9d shows that parameter uncertainty has decreased two times. Therefore, if those ten brightest stars were resolved, their stochastic influence could be minimized with the MCMC model fitting. Deriving parameters of clusters without

brightest stars serves as an illustration of the best cluster parameter accuracy that can be obtained.

Figures 9b and e show model fitting results for the same original images, but here without the three and ten brightest stars, which are located beyond the cluster's FWHM (i.e., in the
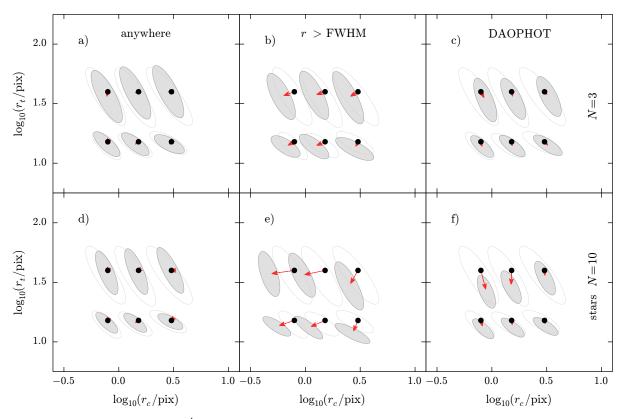
Fig. 9. Star subtraction test on 100 Myr and 10^4 $M_\odot$ King model clusters. 100 models per input parameter node (dot) were computed and 1σ distributions of their recovered parameters are repeated in all panels as open ellipses centered on the nodes. Shaded ellipses indicate distributions of recovered parameters after subtraction of $N = 3$ (*top row*) and 10 brightest stars (*bottom row*) of the cluster. Panels show: **a)** and **d)** results without these stars located anywhere in the cluster; **b)** and **e)** without these stars beyond cluster's FWHM, arrows indicate parameter bias; **c)** and **f)** the performance of FitClust to iteratively fit model and include the 3 and 10 brightest stars detected freely with DAOPHOT.

outskirts), respectively. This test illustrates how parameters are affected when stars located in a cluster's outskirts are preferentially included in model fitting. The size of ellipses indicating parameter uncertainty is smaller than for clusters with all stars; however, core and tidal radii are underestimated.

Since FitClust relies on DAOPHOT to detect stars, it is expected that stars from a cluster's outskirts would be preferentially included in a model fit, as displayed in Fig. 2c, and could result in underestimated tidal radius of the cluster. To test performance of FitClust, we have used the same original images and performed the iterative model fitting procedure described in Sect. 2 by including the three and ten brightest stars detected freely with DAOPHOT. The results are shown in Figs. 9c and f. As in the previous test, the tidal radius of the cluster is underestimated, which illustrates the effect of preferential inclusion of stars in the outskirts. However, DAOPHOT identifies blended groups of stars in a cluster's core and treats them as point sources, therefore the derived core radius stays unaffected, while influence of stochasticity on parameter uncertainty is minimized. To avoid bias of the tidal radius, only a few brightest stars can be considered, as shown in Fig. 9c.

4. Discussion and conclusions

We presented a program, FitClust, developed for automatic derivation of structural parameters of semi-resolved extragalactic star clusters, located in crowded stellar fields. It provides cluster structural parameters and photometry required for evolutionary studies of cluster populations in various galactic environments. The program was tested on real and simulated Subaru Suprime-Cam observations in the M31 galaxy disk.

Recently Brewer et al. (2013) attempted to identify all stars in a simulated frame of a crowded field and to derive their luminosity function by performing a catalog sampling with a variable star number. It took about one day to perform analysis of 1000 stars scattered in a frame of 100×100 pixels in size on a modern multicore PC. As Brewer et al. (2013) suggested, such recovery of crowded fields could also be applicable for analysis of clusters by parameterizing spatial distribution of stars, i.e., derive their structural parameters.

FitClust models cluster as a smooth luminosity distribution with several superposed bright stars and can derive their parameters in a few minutes on multicore PC. Narbutis et al. (2007) studied effects of aperture size on the accuracy of cluster photometry and found that field stars compromise measurements significantly. FitClust provides magnitudes of resolved stars and the unresolved cluster component, which can be used as photometric input data to derive its evolutionary parameters, e.g., de Meulenaer et al. (2013) and Beerman et al. (2012). Photometry of the resolved stars can be used to constrain the age of clusters, once color−magnitude diagrams are constructed and analyzed together with colors of the unresolved cluster component.

Stochastic fluctuations in a cluster's luminosity profile and sky background limit the accuracy of derived structural parameters. While in highly crowded background it is necessary to include background stars in the model fitting to recover parameters of a smooth cluster, in a semi-resolved case stars from a cluster's outskirts would be preferentially detected by DAOPHOT, resulting in underestimated cluster size. Comparing results of star subtraction test, bright star test, and sky background test,

we recommend that the user of FitClust has to set the maximum number of brightest stars to be included in model fit to avoid cluster size underestimation in semi-resolved conditions.

We stress that derived parameter uncertainties are underestimated if, e.g., only the report of MCMC fitting is used. To estimate influence of variable sky background, the image of the cluster can be co-added with a series of background subframes taken close to the cluster and each analyzed separately to build a distribution of derived parameters. Finally, the uncertainty due to stochastic effects, which is dominant for young low mass objects, can be estimated by simulating cluster observations on a star-by-star basis.

In this paper we proposed a method to account for bright field and cluster stars as well as sky background variations on the derived cluster parameters. We have demonstrated a significant influence of stochastic effects which are the main source of structural parameter uncertainty.

The strong points of the implemented algorithm are

– ability to derive cluster structural parameters, which are robust against background variations and bright field stars;
– ability to perform simultaneous photometry of a cluster and resolved bright stars.

In future FitClust versions we will optimize its computing performance and implement consistent treatment of multiband observations.

References

Barmby, P., McLaughlin, D. E., Harris, W. E., Harris, G. L. H., & Forbes, D. A. 2007, AJ, 133, 2764
Bastian, N., Schweizer, F., Goudfrooij, P., Larsen, S. S., & Kissler-Patig, M. 2013, MNRAS, 431, 1252
Beerman, L. C., Johnson, L. C., Fouesneau, M., et al. 2012, ApJ, 760, 104
Brewer, B. J., Foreman-Mackey, D., & Hogg, D. W. 2013, AJ, 146, 7
Deveikis, V., Narbutis, D., Stonkutė, R., Bridžius, A., & Vansevičius, V. 2008, Balt. Astron., 17, 351
Elson, R. A. W., Fall, S. M., & Freeman, K. C. 1987, ApJ, 323, 54
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Goodman, J., & Weare, J., 2010, Comm. App. Math. Comp. Sci., 5, 65
King, I. 1962, AJ, 67, 471
Kodaira, K., Vansevičius, V., Bridzius, A., et al. 2004, PASJ, 56, 1025
Larsen, S. S. 1999, A&AS, 139, 393
Larsen, S. S., de Mink, S. E., Eldridge, J. J., et al. 2011, A&A, 532, A147
de Meulenaer, P., Narbutis, D., Mineikis, T., & Vansevičius, V. 2013, A&A, 550, A20
Narbutis, D., Vansevičius, V., Kodaira, K., Bridžius, A., & Stonkutė, R. 2007, Balt. Astron., 16, 409
Narbutis, D., Vansevičius, V., Kodaira, K., Bridžius, A., & Stonkutė, R. 2008, ApJS, 177, 174
Narbutis, D., Vanagas, R., & Vansevičius, V. 2009, Balt. Astron., 18, 219
Šablevičiūtė, I., Vansevičius, V., Kodaira, K., et al. 2006, Balt. Astron., 15, 547
Šablevičiūtė, I., Vansevičius, V., Kodaira, K., et al. 2007, Balt. Astron., 16, 397
San Roman, I., Sarajedini, A., Holtzman, J. A., & Garnett, D. R. 2012, MNRAS, 426, 2427
Stetson, P. B. 1987, PASP, 99, 191
Stonkutė, R., Vansevičius, V., Arimoto, N., et al. 2008, AJ, 135, 1482
Tanvir, N. R., Mackey, A. D., Ferguson, A. M. N., et al. 2012, MNRAS, 422, 162
Tody, D. 1993, Astronomical Data Analysis Software and Systems II, 52, 173
Vansevičius, V., Kodaira, K., Narbutis, D., et al. 2009, ApJ, 703, 1872
Werchan, F., & Zaritsky, D. 2011, AJ, 142, 48