CHORIZOS

A Bayesian code for the analysis of photometric and spectrophotometric data using parameterized spectral energy distributions

v2.1.4

July 2007

Jesús Maíz Apellániz

Instituto de Astrofísica de Andalucía

1 Introduction

CHORIZOS is a multi-purpose Bayesian code developed to compare photometric data with model spectral energy distributions (SEDs). The user can select the SED family (e.g. Kurucz) and choose the behavior of each parameter (e.g. $T_{\rm eff}$) to be fixed, constrained to a given range, or unconstrained. The code calculates the likelihood for the full specified parameter ranges, thus allowing for the identification of multiple solutions and the evaluation of the full correlation matrix for the derived parameters of a single solution. A descriptión of the code has been published in Maíz Apellániz (2004). CHORIZOS is written in IDL and is distributed to the astronomical comunity in the form of a `.sav` file. The latest version can be obtained from `http://www.iaa.es/~jmaiz` .

The code underwent a major revision to accomodate arbitrary SEDs in v2.0 (December 2005) and is currently fully functional, although a number of upgrades are planned for the future. The following table shows the current and future capabilities:

Characteristic	Current	Future
SED models	User defined Stars: Kurucz, Lejeune, TLUSTY Clusters: Starburst99	Open to suggestions
Number of parameters	≤ 5 total ≤ 3 intrinsic (user defined) ≤ 2 extrinsic (extinction law + reddening)	≤ 6 total ≤ 3 intrinsic (user defined) ≤ 3 extrinsic (user defined)
Parameter control	Full, range, fixed Adjustable grid	Full Bayesian priors
Filters/colors	User-defined filters, 88 preinstalled Colors	Future HST instruments preinstalled Indices, equivalent widths
Spectrophotometry	No	Yes

2 A quick start

2.1 Installing CHORIZOS

Start by creating a directory where the program will be located, making sure that there is enough space there (depending on what SEDs you are interested in, that may mean from several hundreds of Mb to several Gb). We will call the path to that directory `CHORIZOS_DIR` and it should be included in your `IDL_PATH` system variable (see `http://www.linuxhelp.net/forums/lofiversion/index.php/t4853.html` for help on setting up `IDL_PATH`). Go to `http://www.iaa.es/~jmaiz` and download the two files `chorizos.tar.gz` and `kurucz_proc.tar.gz` into `CHORIZOS_DIR`. Note that the two files are quite large (especially the second one), so you will need a fast connection and/or some patience to do this. At the `unix shell >` prompt in `CHORIZOS_DIR` type:

```
unix shell > gunzip chorizos.tar.gz
unix shell > gunzip kurucz_proc.tar.gz
unix shell > tar xfv chorizos.tar
... lots of files ...
unix shell > ln -s chorizos_6.3.sav chorizos.sav
unix shell > tar xfv kurucz_proc.tar
... some files ...
```

And with that you are ready to start using CHORIZOS! But before we do that, here is some additional information:

- CHORIZOS is written in IDL and requires that program to be installed with a license.

- If your version of IDL is earlier than 6.3, substitute `6.3` by `6.0` or by `5.6` (as appropriate) in the `ln` command. IDL versions prior to 5.6 are not supported, sorry.
- You should not need the tar files anymore so if disk space is an issue, go ahead and delete them.
- There are a number of other SED files in the web site which you can also download. More information is provided about them later in this document.

2.2 Executing CHORIZOS

`examples_ubv.dat` is a file included in the `CHORIZOS_PATH/dat` directory distribution that can be used to test how the program works. The data are slightly modified versions of those shown in Figs. 2 and 3 of PASP **116**, 859-875 (2004) and consist of Johnson UBV photometry for nine stars. The file contains the lines necessary to calculate the probability as a function of (a) effective temperature and (b) reddening for $\log g = 5.0$ solar-metallicity standard-Galactic-extinction Kurucz models. Take a look at the file to see the information that it contains: file names to indicate which SED models to use, the selection of filters and colors to be used in the likelihood evaluation, and the data for the nine sample stars.

Start by creating a working directory and copy `examples_ubv.dat` there. Then, in that directory type:

```
unix shell > idl
IDL Version 6.3, Mac OS X (darwin i386 m32). (c) 2006, Research Systems, Inc.
Installation number: 501126.
Licensed for use by: Astronomical Institute of Villaburros de Abajo

IDL> chorizosinit                                        ; We load the CHORIZOS routines
IDL> chorizos, 'examples_ubv.dat', subgrid_n=[9,0,0,9] ; CHORIZOS execution (likelihood evaluation)
... lots of messages and some waiting time ...
IDL> statplots, 'examples_ubv.dat'                       ; STATPLOTS execution (plots and statistics)
... lots of messages and some waiting time ...
```

The output of the the first module (called `CHORIZOS`) should be a file named `examples_ubv.out` located in a newly created directory named `examples_ubv`. Examine it with a text editor to see its contents, which should be similar to the screen output. `CHORIZOS` also generates a file named `examples_ubv.sav` in that same directory and an individual file for each star in the `examples_ubv/objects` subdirectory; those files are used internally by the code to transfer information between modules.

The output of the second module (`STATPLOTS`) is divided into four types:

- A PS file for each star with the most likely (mode) SED, its synthetic photometry, and the observed photometry. These PS files are located in the `examples_ubv/ps` subdirectory and end in `spect.ps`. The associated data are located in the same subdirectory and have the same name with a `dat` suffix instead of a `ps` one.
- A PS file for each star with its probability in the $T_{\rm eff}$-$E(4405-5495)$ plane[1]. These PS files are located in the `examples_ubv/ps` subdirectory and end in `par14.ps`.
- A file named `examples_ubv.parstat` in the `examples_ubv` directory which should contain (in this case) information similar to that in `examples_ubv.out`.
- A file named `examples_ubv.derivstat` in the `examples_ubv` directory which should contain the properties of some derived quantities.

In the following sections we discuss the full capabilities of the CHORIZOS package.

[1] $E(4405-5495)$ is the monochromatic equivalent to $E(B-V)$

3 Basic usage

The CHORIZOS package calculates the probability that a given SED model is consistent with the observed photometry or spectrophotometry using two modules that have to be executed in sequence. The first module, also called `CHORIZOS` (written with a fixed-width font in this document in order to avoid confusion between the names of the package and of the module itself), calculates the likelihood of all the selected SEDs as a function of the N parameters (temperature, reddening, age...). The second module, `STATPLOTS`, applies the Bayesian priors to calculate the final probability as well as the characteristics of the distribution for each parameter and for a number of derived quantities (such as total extinction or age-corrected apparent magnitudes for clusters). `STATPLOTS` also generates a number of plots.

CHORIZOS comes with a series of pre-processed SED models for stars and clusters and with different parameters included (see Appendix A). In this section we explore the basic usage of CHORIZOS, which assumes that the processed SED models are ready. In the following section we describe how to process SEDs in order to tap into the full potentiaal of CHORIZOS.

3.1 The `CHORIZOS` module

3.1.1 Command line syntax

The syntax for executing the second step, CHORIZOS, is the following:

```
PRO CHORIZOS, input_file, $

              (current options)

              SUBGRID_N=subgrid_n, /NO_SAVE, $

              (future options)

              ZP_VNEW=zp_vnew, TEST=test, MOD_S=mod_s, CUBIC_INT=cubic_int
```

`input_file` is the name of the main input file, which should end in `.dat`. See below for more details. Currently, the only option for executing `CHORIZOS` is:

- `SUBGRID_N`: `CHORIZOS` interpolates the magnitudes computed from the SEDs (that previous step is done in `GENSYNPHOT`, see the next section) into a fine grid in order to produce a better sampling of the likelihood . This parameter gives the number of points to add in the fine grid between each of the points in the coarse grid for each of the parameters (i.e. in order to improve the sampling of a given parameter by a factor of two, the corresponding `SUBGRID_N` component should be 1; to improve it by a factor of 4, it should be 3, and so on). `SUBGRID_N` should be a vector with the same number of components as parameters in the processed models (see example in the previous section).
- `/NO_SAVE`: Flag that can be used to cancel variable saving. This option saves time and disk space but, if activated, `STATPLOTS` cannot be executed afterwards.

3.1.2 Input file syntax: magnitude case

I describe the syntax for the input file analyzing the example in the previous section, where the photometry is given by the Johnson UBV magnitudes of the objects. The contents of `examples_ubv.dat` are:

```
kurucz_d_ddd31d.fits           ; Processed model (SED) file
examples_ubv                   ; Output root name
examples_ubv                   ; Output root directory
johnson_u johnson_b johnson_v  ; Input filters
0 1                            ; Color #1 to be used
1 2                            ; Color #2 to be used
star_1 -0.45 0.0264 0.00 0.0264 -0.30 0.0264 NL NL 5.0 5.0 0.0 0.0 NL NL
star_2  0.17 0.0264 0.00 0.0264 -0.43 0.0264 NL NL 5.0 5.0 0.0 0.0 NL NL
star_3  0.74 0.0264 0.00 0.0264 -1.16 0.0264 NL NL 5.0 5.0 0.0 0.0 NL NL
star_4 -0.20 0.0264 0.00 0.0264 -0.99 0.0264 NL NL 5.0 5.0 0.0 0.0 NL NL
star_5 -0.32 0.0264 0.00 0.0264  0.12 0.0264 NL NL 5.0 5.0 0.0 0.0 NL NL
star_6  0.41 0.0264 0.00 0.0264 -0.49 0.0264 NL NL 5.0 5.0 0.0 0.0 NL NL
star_7  0.77 0.0264 0.00 0.0264 -0.92 0.0264 NL NL 5.0 5.0 0.0 0.0 NL NL
star_8 -0.06 0.0264 0.00 0.0264 -0.50 0.0264 NL NL 5.0 5.0 0.0 0.0 NL NL
star_9  0.30 0.0264 0.00 0.0264 -0.72 0.0264 NL NL 5.0 5.0 0.0 0.0 NL NL
```

- The first line gives the name of the processed models to be used. The two files (with the same name) are in the directories `CHORIZOS_DIR/seds` and `CHORIZOS_DIR/magtables`. Note that everything past the semicolon is interpreted as a comment (but do not use tabs here or anywhere else since IDL interprets them differently than white spaces). In this case, the models we are using have $N = 4$ parameters: effective temperature, gravity, metallicity, and reddening.
- The second and third line give the root name and directory for the output, which will be used by `CHORIZOS` and by the following module, `STATPLOTS`. Here they are the same but the user can make them different.
- The fourth line lists the F filters used. The default choices for filters are listed in Appendix B. The order is used to establish the filter numbering system for the input data and for the output. Following the IDL convention, the numbering starts with zero.
- The next group of lines lists the M (two in this case) colors used for the likelihood calculation according to the order defined in the previous line (here, Johnson $U - B$ and $B - V$).
- The last group of lines (one per object) contains the following information:
  - The first column is the object name. No spaces are allowed since they are used to separate the subsequent columns.
  - The next $2 \times F$ columns are the magnitudes and uncertainties for each filter in the order m_0, σ_0, m_1, $\sigma_1 \ldots m_{F-1}$, σ_{F-1}. All input magnitudes and colors in CHORIZOS are given in the modified Vega magnitude system (VEGAMAG with non-zero ZP_P for some filters), as described in Maíz Apellániz (2006) and Appendix B.
  - The next $2 \times N$ columns list the minimum and maximum values allowed for each of the N parameters in the order min_1, $\mathrm{max}_1 \ldots \mathrm{min}_N$, max_N. `NL` indicates the limit is not used and the full range of the parameter is used instead (see Appendix A). In this example we are leaving the first and last parameters (T_{eff}, $E(4405 - 5495)$) unconstrained and fixing $\log g$ to be 5.0 and the metallicity to be solar. Therefore, the number of degrees of freedom, N_{free}, for each star is two.

3.1.3 Input file syntax: color case

The above example assumes that the available photometry is provided as magnitudes. Suppose instead that the photometric information you have for your objects consists of colors. CHORIZOS can also handle that case and another file in the distribution, `CHORIZOS_DIR/dat/examples_ubv2.dat`, illustrates how to proceed in that case:

```
kurucz_d_ddd31d.fits                    ; Model file
examples_ubv2                           ; Output root name
examples_ubv2                           ; Output root directory
johnson_u-johnson_b johnson_b-johnson_v ; Input colors
0                                       ; Color #1 to be used
1                                       ; Color #2 to be used
star_1 -0.45 0.0373  0.30 0.0373  NL   NL   5.0  5.0  0.0  0.0  NL   NL
star_2  0.17 0.0373  0.43 0.0373  NL   NL   5.0  5.0  0.0  0.0  NL   NL
star_3  0.74 0.0373  1.16 0.0373  NL   NL   5.0  5.0  0.0  0.0  NL   NL
star_4 -0.20 0.0373  0.99 0.0373  NL   NL   5.0  5.0  0.0  0.0  NL   NL
star_5 -0.32 0.0373 -0.12 0.0373  NL   NL   5.0  5.0  0.0  0.0  NL   NL
star_6  0.41 0.0373  0.49 0.0373  NL   NL   5.0  5.0  0.0  0.0  NL   NL
star_7  0.77 0.0373  0.92 0.0373  NL   NL   5.0  5.0  0.0  0.0  NL   NL
star_8 -0.06 0.0373  0.50 0.0373  NL   NL   5.0  5.0  0.0  0.0  NL   NL
star_9  0.30 0.0373  0.72 0.0373  NL   NL   5.0  5.0  0.0  0.0  NL   NL
```

- This example provides the same colors as the previous one. Note, however, that if you execute this example, the CHORIZOS output will be similar to the previous one but not identical because the colors here are not correlated among them (i.e. they are independent quantities).
- The fourth line now lists the C colors in the photometry and the next M lines the ones actually used in the calculation (it is possible to have e.g. four colors per object and use only three simply by changing these lines).
- The last group of lines is similar to the previous case but here colors are listed instead of magnitudes according to c_0, σ_0, c_1, $\sigma_1 \ldots c_{C-1}$, σ_{C-1}.

3.1.4 CHORIZOS output

CHORIZOS generates both a screen output and a file in the output directory. The information content is similar; here we describe the file contents.

- The first two columns are the object number and name.
- The next $8 \times N$ columns are, for each parameter in sequence: [a] a flag that indicates whether the parameter was left free (1) or was fixed (0), [b] its mean, [c] its median, [d] its mode (calculated in an ultrafine grid with a $10\times$ sampling, which is activated only if not located at one of the grid edges; otherwise, the regular grid is used), [e] its standard deviation, [f] its lower-standard-deviation-equivalent measured from the percentile values assuming a Gaussian distribution, [g] its upper-standard-deviation-equivalent measured from the percentile values assuming a Gaussian distribution, and [h] a grid-sampling test that measures whether the value used for SUBGRID_N for the parameter is appropriate or no for the provided uncertainties (a value > 1.0 is unacceptable while one < 0.1 is good; the user should decide what to do if it is in between those two values on a case by case basis).
- The next column is a flag that should be zero if the mode was found at one of the grid edges, non-zero otherwise.
- The next two columns give the minimum value of χ^2 (NOT divided by $N_{\rm free}$) in the fine and ultrafine grids, respectively. If the mode is located at one of the grid edges, a dummy value is used for the second value.
- The next column gives $N_{\rm free}$ for the object.
- The last $3 \times M$ columns are, for each color in sequence: [a] the measured uncertainty, [b] the difference between the model mode value (in the fine grid) and the measured color, and [c] the ratio of the second of those quantities divided by the first one (i.e. the normalized distance between the measured color and the best model one). When the grid-sampling test for all parameters is $\ll 1.0$, the last column should be ~ 1.0 if $M > N$ and $\ll 1.0$ if

$M = N$ for all colors. If that is not the case, it is an indication that there is a problem with that color for that object (due e.g. to an incorrect data value or to the fact that the object cannot be described by the selected SEDs).

3.2 The STATPLOTS module

3.2.1 Command line syntax

The syntax for executing the second module, STATPLOTS, is the following:

```
PRO STATPLOTS, input_file, $

               (current options)

               OBJ_I=obj_i, NO_SP=no_sp, NO_CONT=no_cont, NO_STAT=no_stat, $
               PAR_LIM=par_lim, SP_LIM=sp_lim, SP_MEAN=sp_mean, $
               C_MIN=c_min, C_MAX=c_max, FILT_REF=filt_ref, MONT_N=mont_n, $
               EXTRA_SP=extra_sp, EXTRA_CONT=extra_cont, $

               (future options)

               FINE_CONT=fine_cont
```

input_file is the same file as before. The options are:

- OBJ_I: Vector of indices to select which objects to analyze. By default, all objects are analyzed.
- /NO_SP: Flag not to do spectrum (SED) plots.
- /NO_CONT: Flag not to do probability contour plots.
- /NO_STAT: Flag not to calculate statistics.
- PAR_LIM: Vector with the parameter limits following the order $\min_1$, $\max_1 \ldots \min_N$, $\max_N$. If not used, the same limits as for the previous module are used instead.
- SP_LIM: Limits for spectrum plots in the order $[1/\lambda_{\max}, m_{\mathrm{AB},\max}, 1/\lambda_{\min}, m_{\mathrm{AB},\min}]$.
- /SP_MEAN: Flag to select the mean values of the parameters instead of the mode for the spectrum plots.
- C_MIN: Minimum contour to be displayed in the probability plots normalized to the maximum value in the plot (by default it is set to 0.05).
- C_MAX: Maximum contour to be displayed in the probability plots normalized to the maximum (by default it is set to 0.95).
- FILT_REF: Name of the reference filter, which overrides the automatic selection. The name should be given inside quotation marks. This filter is the one used to calculate the extinction, bolometric corrections, and associated quantities.
- MONT_N: Number of Montecarlo samples for the analysis of the correlation between the magnitude in the reference filter and the rest of the quantities. The default value is 10. Increasing it should arise the reliability of the output but at the possible cost of considerable memory use and processing time.

- EXTRA_SP: Additional parameters for the SINGLEPLOT command in the spectrum plots. See the documentation for the package JMAPLOT at http://www.iaa.es/~jmaiz for further details. All SINGLEPLOT options are allowed. An example would be: EXTRA_SP="COLOR=[1,2,4], TITLE='My spectrum plot'".
- EXTRA_CONT: Additional parameters for the SINGLECONTOUR command in the contour plots. See the documentation for the package JMAPLOT at http://www.iaa.es/~jmaiz for further details. All SINGLECONTOUR options are allowed except EXTERNAL. An example would be: EXTRA_CONT="AXISF=2., DAXIS=[1,1,0,0], TITLE=''".

3.2.2 STATPLOTS output: spectrum plots

A spectrum plot is created for each star in the ps subdirectory of the output root directory with a file name that ends in spect.ps. The associated data are located in the same subdirectory and have the same name with a .dat suffix instead of a .ps one. The plot includes the "best" SED (mode or mean[2]), the associated synthetic photometry (indicated by star symbols), and the observed photometry (indicated by the points with the error bars, the horizontal ones showing the approximate extent of the filter and the vertical ones the photometric uncertainties). The horizontal scale is shown as inverse wavelength (in μ^{-1}) and the vertical scale as AB magnitudes.

If input_file contains colors and not magnitudes, then the observed photometry is not plotted and an artificial zero point is added to the other values for normalization purposes. This can be overridden by creating a file with the same root as input_file but substituting the .dat suffix by .mag and listing the magnitudes and uncertainties there. See the file CHORIZOS_DIR/examples_ubv2.mag for an example.

3.2.3 STATPLOTS output: probability contour plots

STATPLOTS multiplies the likelihood calculated by CHORIZOS by a Bayesian prior which is a function of the model parameters[3] to obtain an $N_{\rm free}-$dimensional probability distribution. If $N_{\rm free} > 2$, such a distribution is hard (or impossible) to visualize directly. Therefore, STATPLOTS collapses the distribution into every possible pair of parameters and displays the result as a probability contour plot. The contour plots are written in the ps subdirectory of the output root directory and end in parij.ps, where $i = 1, N-1$ and $j = i+1, N$. Note that plots where i or j correspond to a fixed parameter are not derived. The values of the parameters of the SED selected for the spectrum plot are indicated with a white star symbol.

3.2.4 STATPLOTS output: statistics

STATPLOTS generates two statistical output files. The first one has a .parstat suffix, is located in the output root directory, and contains an information similar to the CHORIZOS output but here the results are calculated after multiplying the likelihood by the Bayesian prior (note, however, that in the current version of the package no new mode calculation is performed; the result is simply copied from the previous step).

The second file has a .derivstat suffix and is also located in the output root directory. It always contains the following quantities: the extinction A_r in the reference filter r, the bolometric correction in the same filter BC_r, and $A_r - BC_r$. In each case the mean and standard deviations are provided. Some additional quantities are also provided if input_file contains magnitudes:

- The extinction-corrected apparent magnitude in the reference filter $m_r - A_r$, where m_r is the measured magnitude in the reference filter, and the extinction-corrected apparent bolometric magnitude $m_r + BC_r - A_r$.
- For cluster models, the zero-age extinction-corrected apparent magnitude in the reference filter $m_r - C_r - A_r$, where the age-correcton factor C_r is introduced to take into account age-dimming effects.

[2]Note that in the current version of the package, the mode is not recalculated by STATPLOTS while the mean is.

[3]The current version of the package allows only for the use of step- and box- functions for the prior. Future versions will include arbitrary functions.

- For temperature-dependent stellar models, the covariance and correlation between $T_{\rm eff}$ and $m_r + BC_r - A_r$, which are needed to generate the uncertainty ellipses in a theoretical ($T_{\rm eff}$ vs. $M_{\rm bol}$) HR diagram.

Quantities that involve m_r and the model parameters are computed by means of a Montecarlo simulation.

4 Advanced usage

The full power of CHORIZOS appears when the user is able to control the SED families, parameters, and filters. In this section I describe how to do that.

4.1 The `GENSYNPHOT` module

4.1.1 Description and execution syntax

`GENSYNPHOT` is a module that can be executed before `CHORIZOS` that allows the user to "process" an "original" SED family. An original family is a collection of SEDs with a wide range of up to three "intrinsic" (i.e. depending on the characteristics of the object itself) parameters. The information about the original SED family is located in a single FITS table in the directory `CHORIZOS_PATH/seds` (see Appendix A). For example, `kurucz_orig.fits` contains the SEDs for Kurucz models with $T_{\rm eff}$ between 3 500 and 50 000 K, $\log g$ between 0.0 and 5.0 cgs, and $\log Z$ between -1.5 and 0.0 in solar units.

The processing done by `GENSYNPHOT` takes place in three steps:

1. The original grid is resampled into a new grid. An intrinsic parameter in the original SED family can be left as a functioning parameter in the processed family (with a range of values); it can be fixed to a given value (e.g. only a $T_{\rm eff}$ of 30 000 K may be included); or it can be left as a "dead" parameter, by making it a function of one or several of the others (an example would be gravity for main-sequence models).
2. The grid is then possibly extended using "extrinsic" (i.e. depending on characteristics external to the objects themselves) parameters. Currently, the two available options are the extinction law and the amount of reddening.
3. Finally, the synthetic photometry of the full (intrinsic + extrinsic) grid is calculated. Note that the photometry is performed directly on the reddened SEDs (i.e. no Q-factor approximations are involved).

In principle, one could use `GENSYNPHOT` just once for an original SED family using all possible intrinsic and extrinsic parameters and all filters (the processed file `kurucz_d_ddd31dd.fits` does almost that, but uses a fixed extinction law) but there are several reasons why that is not practical:

- Such an execution would be very long (weeks) and the generated files extremely large (at least several Gb). In some instances it makes more sense to reduce the parameter range or the filter selection, tailoring them for a specific application.
- New filters may need to be added or the SED themselves modified, thus requiring a new execution of the module. This is dealt with in the next two subsubsections.
- For some applications, it makes sense to use dead parameters to eliminate unnecessary degrees of freedom. For example, one can fix the gravity of Kurucz models as a function of temperature to represent supergiant stars if the objects analyzed are known to be of that luminosity class.

The syntax for executing `GENSYNPHOT` is:

```
PRO GENSYNPHOT, $

                (current options)

                MOD_FILE=mod_file
```

The single option for executing `GENSYNPHOT` is:

- `MOD_FILE`: The name of the file that contains the information on which original SED family to use and how to process it. Its syntax is described in the next subsubsection. The file should be located in the directory `CHORIZOS_DIR/dat` and will be referred to as the model data file. The default value is `'kurucz_d_ddd31d.dat'`.

The `GENSYNPHOT` output consists of two FITS tables, each one with the same file name, one in the directory `CHORIZOS_DIR/seds` and the other one in `CHORIZOS_DIR/magtables`. The first one contains the processed SEDs in the intrinsic-parameter processed subgrid and the second one the synthetic photometry in the full processed grid.

4.1.2 Model data file syntax

I show what the syntax for the model data file should be using an example contained in the CHORIZOS distribution. The first file is `kurucz_d_ddd31d.dat`:

```
kurucz_orig.fits       ; Input (original) models
kurucz_d_ddd31d.fits   ; Output (processed) models
def_filt_syst.dat      ; Default filter systems
1 default              ; First  original parameter (Temperature     ) is variable and uses default scale
1 default              ; Second original parameter (log(Gravity)    ) is variable and uses default scale
1 default              ; Third  original parameter (log(Metallicity)) is variable and uses default scale
0 3.1                  ; R_5495                                       is fixed    and equal to 3.1
1 default              ; E(4405-5495)                                 is variable and uses default scale
```

- The first two lines give the names for the original and processed SED families.
- The third line gives the file name with the information on which filter systems to use. The file must be located in `CHORIZOS_DIR/dat` and its syntax is described in the next subsection. `def_filt_syst.dat` selects the filters in Appendix A (a different file, `johnson.dat` is provided in the CHORIZOS distribution as an alternative example).
- The next three lines describe how to process the original parameters. The first column should be a 1 if the user wants the parameter to be left functioning or 0 if he/she wants it to be fixed or dead. The second column describes the values themselves.
  - For a functioning parameter, this can be `default` (which selects the same range as the original parameter) as in the example shown, or any valid IDL function (for example, `10000.0+5000*FINDGEN(3)` for the temperature would generate the three values 10 000, 15 000, and 20 000) or vector (e.g. `[10000.,15000.,20000.]` would be equivalent to the previous example). No spaces in the text are allowed in either case.
  - For a fixed parameter, the value should be given (e.g. `5.0` or `alog10(2)`).
  - For a dead parameter, a function should be given in which the original parameters are referred to as `v1`, `v2`, and `v3`, respectively. For example, using `-0.1*v2` for the metallicity would select values equal to minus one tenth of the gravity. A more practical example would be `main_sequence_kurucz(v1)` for the gravity (that function is included in the CHORIZOS distribution, the user is encouraged to take a look at it).
- The next two lines are the equivalent of the previous three for the two extrinsic parameters:

– The first one defines the extinction law, parameterized by R_{5495}, the monochromatic equivalent of R_V. CHORIZOS includes the Galactic extinction laws of Cardelli et al. (1989), the average LMC and LMC2 laws of Misselt et al. (1999), and the SMC law of Gordon and Clayton (1998). The value of R_{5495} in CHORIZOS is defined to be the true one for the Galactic laws while the Magellanic Cloud extinctions are assigned artificial values of 1.0, 0.0, and −1.0 for the average LMC, LMC2, and SMC respectively[4]. The default value is equivalent to the IDL function `-1.0+0.1*FINDGEN(71)` (all laws with a $\Delta R_{5495} = 0.1$).

– The second one defines the reddening, parameterized by $E(4405-5495)$, the monochromatic equivalent of $E(B-V)$. The default value is equivalent to the IDL function `-0.5+0.1*FINDGEN(56)` (all values between −0.5 and 5.0 with a $\Delta E(4405-5495) = 0.1$) .

4.2 Adding new filters

As shown in Appendix B, filters in CHORIZOS are grouped in sets and their names are defined using the convention `set_filter` (e.g. `jonhson_u` is a filter named `u` which belongs to the `johnson` set). New sets can be easily added, as long as their names do not conflict with the existing ones. Suppose we want to create a new set named `myset` with three filters named `a`, `b`, and `c`. The following steps should be followed:

- In the directory `CHORIZOS_PATH/throughputs` create three ASCII files named `myset_a.dat`, `myset_b.dat`, and `myset_c.dat`, with two columns each, one for the wavelength and one for the throughput. The wavelength should be in Angstroms and the throughput in photon-counting form (Maíz Apellániz 2006).

- Create a file named `myset.dat` in the directory `CHORIZOS_PATH/dat` similar to the one below (you are welcome to take a look at `def_filt_syst.dat` in the same directory to see possible variations). Each set is defined with two lines. The first one contains the name of the set and of each filter. The second one contains the index of the default reference filter for the set (in this case below, 1, which corresponds to `myset_b`) and the zero points for each filter, defined as in Eqn. 1. As with other CHORIZOS input files, spaces should be used instead of tabs. The reader is referred to Maíz Apellániz (2006) for a discussion on zero points.

```
;          0       1       2  Any line that starts with a semicolon is ignored
myset      a       b       c
1     +2.718 -3.1416   0.000
```

- In IDL, execute the command: `IDL> bands, FILT_FILE='myset.dat'`. The `BANDS` module in CHORIZOS computes the basic properties of each filter and resamples the throughputs in the wavelength grid used by all modules (which covers the 1 000 - 10 000 Å range in 1 Å steps and the 10 000 - 25 000 range in 10 Å steps). The `BANDS` output is a modified version of `filt_info.dat` in `CHORIZOS_PATH/dat` and three new files named `myset_a_res.dat`, `myset_b_res.dat`, and `myset_c_res.dat` in `CHORIZOS_PATH/throughputs`.

- In IDL, execute the command: `IDL> vega_stab, FILT_FILE='myset.dat'`. The `VEGA_STAB` module in CHORIZOS computes the transformations between the modified Vega magnitude system (VEGAMAG with non-zero ZP_P for some filters) and the ST and AB magnitude systems.

- Now, you are ready to execute `GENSYNPHOT`. If you are interested in `myset` magnitudes only, simply substitute `def_filt_syst` by `myset` in the third line of the model data file (see previous subsection). If you want to combine different filter sets, you should create a new file in `CHORIZOS_PATH/dat` with the appropriate selections from `def_filt_syst.dat` and `myset.dat`.

[4]Those values are artificial because the true R_{5495} for the Magellanic Cloud extinction laws are actually close to 3.1. The definition is done in order to provide a simple parameterization in the extinction law.

4.3 Generating SED original families

An original SED family is described in a FITS tables that contain one vector with the wavelengths (with `wav_n` elements), one to three vectors with the values of the intirnsic parameters (with `mod_n` elements each), and one array with the SEDs (in units of erg s^{-1} cm^{-2} $Å^{-1}$ and with dimensions of `mod_n` × `wav_n`). The wavelength vector can be of any length, since `GENSYNPHOT` automatically resamples it. The models have to be arranged in a convex Cartesian grid with no holes in it (e.g. the standard Kurucz or TLUSTY grids are OK).

The header of the FITS table has to contain information about the models in order for `GENSYNPHOT` to process them. If you are interested in learning how to create them, probably the best way to start is with the `prokurucz.pro` and `prosb99.pro` files included in the distribution. Those were the IDL procedures used to create `kurucz_orig.fits` and `sb99_def_orig.fits` from the files obtained from the Kurucz (`http://kurucz.harvard.edu`) and Starburst99 (`http://www.stsci.edu/science/starburst99`) web pages.

4.4 Obtaining information about the FITS tables

In order to see the contents of the FITS tables (either for the original or processed SEDs), one can use the following IDL instructions (the example uses `kurucz_orig.fits` but is applicable to all FITS files in Appendix A):

```
IDL> path = chorizospath()
IDL> v    = mrdfits(path + '/seds/kurucz_orig.fits',1,hdr)
MRDFITS: Binary table.  5 columns by  1 rows.
IDL> help, v, /str
** Structure <306e210>, 5 tags, length=8004948, data length=8004948, refs=1:
   WAV             FLOAT     Array[1221]
   V1              FLOAT     Array[1634]
   V2              FLOAT     Array[1634]
   V3              FLOAT     Array[1634]
   FLAM            FLOAT     Array[1634, 1221]
IDL> print, hdr
XTENSION= 'BINTABLE'           /Binary table written by MWRFITS v1.4a
BITPIX  =                    8 /Required value
NAXIS   =                    2 /Required value
NAXIS1  =              8004948 /Number of bytes per row
NAXIS2  =                    1 /Number of rows
PCOUNT  =                    0 /Normally 0 (no varying arrays)
GCOUNT  =                    1 /Required value
TFIELDS =                    5 /Number of columns in table
COMMENT
COMMENT  *** End of mandatory fields ***
COMMENT
NAME    = 'Kurucz models'           /
MODTYPE = 'Star'                    /
MODN    =  1634                    /
PARN    =     3                     /
WAVN    =  1221                     /
WAVUNIT = 'Angstrom'                /
PARNAME1= 'T'                       /
PARNAML1= 'Temperature'             /
PARUNIT1= 'K'                       /
PARLOG1 =     0                     /
PARMIN1 =  3500                     /
```

```
PARMAX1 = 50000                    /
PARNAME2= 'g'                      /
PARNAML2= 'Gravity'                /
PARUNIT2= 'cgs'                    /
PARLOG2 =     1                    /
PARMIN2 =  0.00                    /
PARMAX2 =  5.00                    /
PARNAME3= 'Z'                      /
PARNAML3= 'Metallicity'            /
PARUNIT3= 'Solar units'            /
PARLOG3 =     1                    /
PARMIN3 = -1.50                    /
PARMAX3 =  0.00                    /
COMMENT
COMMENT  *** Column names ***
COMMENT
TTYPE1  = 'WAV     '           /
TTYPE2  = 'V1      '           /
TTYPE3  = 'V2      '           /
TTYPE4  = 'V3      '           /
TTYPE5  = 'FLAM    '           /
COMMENT
COMMENT  *** Column formats ***
COMMENT
TFORM1  = '1221E   '           /
TFORM2  = '1634E   '           /
TFORM3  = '1634E   '           /
TFORM4  = '1634E   '           /
TFORM5  = '1995114E'           /
COMMENT
COMMENT  *** Column dimensions (2 D or greater) ***
COMMENT
TDIM5   = '( 1634, 1221)'      /
COMMENT Kurucz models for Z=-1.50 SDSC GRID  [-1.5]   VTURB 2.0 KM/S    L/H 1.25
COMMENT                   Z=-1.00 SDSC GRID  [-1.0]   VTURB 2.0 KM/S    L/H 1.25
COMMENT                   Z=-0.50 SDSC GRID  [-0.5]   VTURB 2.0 KM/S    L/H 1.25
COMMENT                   Z= 0.00 SDSC GRID  [+0.0]   VTURB 2.0 KM/S    L/H 1.25
END
```

Alternatively, CHORIZOS includes `MODELINFO`, a module that prints a summary of the information above:

```
IDL> modelinfo,  'kurucz_orig.fits'
Using MODELINFO v2.1.4

1 model(s) found for kurucz_orig.fits

1/1 ****************************************
seds/kurucz_orig.fits is a file that contains original models
There are a total of 1634 Kurucz models

3 intrinsic parameter(s):
#1 is Temperature measured in K between 3500 and 50000
```

```
#2 is log(Gravity) measured in cgs between 0.00000 and 5.00000
#3 is log(Metallicity) measured in Solar units between -1.50000 and 0.00000
```

`MODELINFO` allows for the use of wildcards (try e.g. `IDL> modelinfo, 'kurucz*.fits'`). Also, if the `/COMMENT` flag is added, a more verbose output is obtained.

5 Some (not-so-famous, which is not the same as infamous) last words

IDL is a very powerful language but it has its limitations, most notoriously its speed and memory usage. Do not be surprised if more than once you find yourself staring at a screen for days waiting for `GENSYNPHOT` to process an SED family or having `CHORIZOS` or `STATPLOTS` complain about not having enough memory. Here are some tips on what to do when such a problem arises.

- When executing `GENSYNPHOT`, use only the filters you are interested in and do not include free parameters you are not going to need or extend them beyond the range of physical values that are possible for your datasets.
- In `CHORIZOS` a good approach is to do an initial run without using `SUBGRID_N` to test the range of possible values, constrain the parameters accordingly in `input_file`, and then activate `SUBGRID_N` in a second run.
- In `STATPLOTS` typically the main speed and memory drain is the statistical calculation. If you are only interested in the plots, use `NO_STAT` or set `MONT_N` to a number smaller than the default.

I would like to thank all the people who have helped me test CHORIZOS, especially Leonardo Úbeda.

Bibliography

Cardelli, J. A., Clayton, G. C., and Mathis, J. S. 1989, *ApJ* **345**, 245

Cohen, M., Wheaton, W. A., and Megeath, S. T. 2003, *AJ* **126**, 1090

Gordon, K. D. and Clayton, G. C. 1998, *ApJ* **500**, 816

Holberg, J. B. and Bergeron, P. 2006, *AJ* **132**, 1221

Laidler, V. et al. 2005, *Synphot User's Guide v5.0 (STScI: Baltimore)*

Maíz Apellániz, J. 2004, *PASP* **116**, 859

Maíz Apellániz, J. 2006, *AJ* **131**, 1184

Maíz Apellániz, J. 2007, in C. Sterken (ed.), *The Future of Photometric, Spectrophotometric and Polarimetric Standardization*, Vol. 364 of *ASP Conf. Series*, 227

Misselt, K. A., Clayton, G. C., and Gordon, K. D. 1999, *ApJ* **515**, 128

Appendix A: Available SED families

The following is the list of currently available SEDs. Both original (to be used in `GENSYNPHOT`) and processed (to be used in `CHORIZOS`) SEDs are shown. For the original SEDs only the intrinsic parameters are listed, since the extrinsic parameters are added in `GENSYNPHOT`. All the processed SEDs listed here use the default filter set (see Appendix B) but the user can generate his/her own with only a subsample of them (see subsection 4.2). The information for the original SEDs is located in a single file in the directory `CHORIZOS_DIR/seds` while that for the processed SEDs is located in two files, each one with the same name, located in `CHORIZOS_DIR/seds` and `CHORIZOS_DIR/magtables`, respectively.

File name	SED type	Object type	Used parameter(s)	Fixed parameter(s)
`kurucz_orig.fits`	Original	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K $\log(g/1$ cgs$)$= 0.0 - 5.0 $\log(Z/Z_{\odot})$= −1.5 - 0.0	
`kurucz_d_ddd31d.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K $\log(g/1$ cgs$)$= 0.0 - 5.0 $\log(Z/Z_{\odot})$= −1.5 - 0.0 $E(4405-5495)$= −0.5 - 5.0	R_{5495}= 3.1
`kurucz_d_dms003100.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K	Main sequence gravity $\log(Z/Z_{\odot})$= 0.0 R_{5495}= 3.1 $E(4405-5495)$= 0.0
`kurucz_d_dsg003100.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K	Supergiant gravity $\log(Z/Z_{\odot})$= 0.0 R_{5495}= 3.1 $E(4405-5495)$= 0.0
`kurucz_d_dms00dd.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K R_{5495}= 2.0 - 6.0 + MC laws $E(4405-5495)$= −0.5 - 5.0	Main sequence gravity $\log(Z/Z_{\odot})$= 0.0
`kurucz_d_dmsm05dd.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K R_{5495}= 2.0 - 6.0 + MC laws $E(4405-5495)$= −0.5 - 5.0	Main sequence gravity $\log(Z/Z_{\odot})$= −0.5
`kurucz_d_dmsm10dd.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K R_{5495}= 2.0 - 6.0 + MC laws $E(4405-5495)$= −0.5 - 5.0	Main sequence gravity $\log(Z/Z_{\odot})$= −1.0
`kurucz_d_dsg00dd.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K R_{5495}= 2.0 - 6.0 + MC laws $E(4405-5495)$= −0.5 - 5.0	Supergiant gravity $\log(Z/Z_{\odot})$= 0.0
`kurucz_d_dsgm05dd.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K R_{5495}= 2.0 - 6.0 + MC laws $E(4405-5495)$= −0.5 - 5.0	Supergiant gravity $\log(Z/Z_{\odot})$= −0.5
`kurucz_d_dsgm10dd.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K R_{5495}= 2.0 - 6.0 + MC laws $E(4405-5495)$= −0.5 - 5.0	Supergiant gravity $\log(Z/Z_{\odot})$= −1.0
`lejeune_orig.fits`	Original	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K $\log(g/1$ cgs$)$= 0.0 - 5.0 $\log(Z/Z_{\odot})$= −1.5 - 0.0	
`lejeune_d_ddd31d.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K $\log(g/1$ cgs$)$= 0.0 - 5.0 $\log(Z/Z_{\odot})$= −1.5 - 0.0 $E(4405-5495)$= −0.5 - 5.0	R_{5495}= 3.1

File name	SED type	Object type	Used parameter(s)	Fixed parameter(s)
`lejeune_d_dms003100.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K	Main sequence gravity
				$\log(Z/Z_\odot)$= 0.0
				R_{5495}= 3.1
				$E(4405-5495)$= 0.0
`lejeune_d_dsg003100.fits`	Processed	Stars	$T_{\rm eff}$ = 3 500 - 50 000 K	Supergiant gravity
				$\log(Z/Z_\odot)$= 0.0
				R_{5495}= 3.1
				$E(4405-5495)$= 0.0
`sb99_def_orig.fits`	Original	Clusters	log(age/1 a)= 6.0 - 10.0	
			$\log(Z)$= 0.004 - 0.020	
`sb99_def_d_ddGald.fits`	Processed	Clusters	log(age/1 a)= 6.0 - 10.0	
			$\log(Z)$= 0.004 - 0.020	
			R_{5495}= 2.0 - 6.0	
			$E(4405-5495)$= −0.5 - 5.0	
`tlusty_orig.fits`	Original	Stars	$T_{\rm eff}$ = 27 500 - 55 000 K	
			$\log(g/1$ cgs)= 3.00 - 4.75	
			$Z/Z_\odot$ = 0.2 - 1.0	
`tlusty_d_ddd31d.fits`	Processed	Stars	$T_{\rm eff}$ = 27 500 - 55 000 K	R_{5495}= 3.1
			$\log(g/1$ cgs)= 3.00 - 4.75	
			$Z/Z_\odot$ = 0.2 - 1.0	
			$E(4405-5495)$= −0.5 - 5.0	

Appendix B: Preinstalled filters

The list in the next page gives the preinstalled filters, which are defined in the file `def_filt_syst.dat` in the directory `CHORIZOS_DIR/dat`. Those in boldface type are the default reference filters in each set. For each filter, the name has two parts: that of the set (Johnson, Cousins, HST/WFCP2, HST/STIS/CCD, HST/STIS/NUV-MAMA, HST/STIS/FUV-MAMA, HST/ACS/WFC, HST/ACS/HRC, HST/ACS/SBC, HST/NICMOS, Galex, Strömgren, Tycho, SDSS, and 2MASS) and that specific to each filter, with the two parts separated by a point. For each filter P we show the zero point ZP_P used to calculate the magnitude m_P in the modified Vega magnitude system according to:

$$m_P = -2.5 \log_{10} \left(\frac{\int P(\lambda) f_\lambda(\lambda) \lambda \, d\lambda}{\int P(\lambda) f_{\lambda,\mathrm{Vega}}(\lambda) \lambda \, d\lambda} \right) + \mathrm{ZP}_P, \quad (1)$$

where $P(\lambda)$ is the total-system sensitivity curve, $f_\lambda(\lambda)$ is the SED of the object, and $f_{\lambda,\mathrm{Vega}}(\lambda)$ is the SED of the reference spectrum (Vega). The values for ZP_P are taken from Cohen et al. (2003), Maíz Apellániz (2006), and Holberg and Bergeron (2006).

Note that this modified Vega magnitude system differs from the standard VEGAMAG system (Laidler et al. 2005) in the addition of non-zero values of ZP_P for some filters (Maíz Apellániz 2006). The system used by CHORIZOS allows for published Johnson, Cousins, VEGAMAG HST magnitudes, Strömgren, Tycho, and 2MASS magnitudes to be directly used. GALEX and SDSS published magnitudes, however, are given in the AB system and must be converted into a Vega-based system before they are used in CHORIZOS. More details are given in Maíz Apellániz (2007)

#	Filter name	ZP_P	#	Filter name	ZP_P	#	Filter name	ZP_P
1	johnson_u	0.055	31	stisnuv_f25qtz	0.000	61	**acssbc_f115lp**	0.000
2	johnson_b	0.034	32	stisnuv_f25cn182	0.000	62	acssbc_f122m	0.000
3	**johnson_v**	0.026	33	stisnuv_f25ciii	0.000	63	acssbc_f125lp	0.000
4	cousins_r	0.030	34	stisnuv_f25cn270	0.000	64	acssbc_f140lp	0.000
5	**cousins_i**	0.017	35	stisnuv_f25mgii	0.000	65	acssbc_f150lp	0.000
6	wfpc2_f170w	0.000	36	stisfuv_f25lya	0.000	66	acssbc_f165lp	0.000
7	wfpc2_f255w	0.000	37	**stisfuv_25mama**	0.000	67	**nicmos_f110w**	0.000
8	wfpc2_f300w	0.000	38	stisfuv_f25srf2	0.000	68	nicmos_f160w	0.000
9	wfpc2_f336w	0.000	39	stisfuv_f25qtz	0.000	69	nicmos_f187w	0.000
10	wfpc2_f380w	0.000	40	acswfc_f435w	0.000	70	nicmos_f190n	0.000
11	wfpc2_f439w	0.000	41	acswfc_f475w	0.000	71	nicmos_f205w	0.000
12	wfpc2_f450w	0.000	42	acswfc_f550m	0.000	72	nicmos_f222m	0.000
13	wfpc2_f467m	0.000	43	**acswfc_f555w**	0.000	73	galex_fuv	0.000
14	wfpc2_f502n	0.000	44	acswfc_f606w	0.000	74	**galex_nuv**	0.000
15	wfpc2_f547m	0.000	45	acswfc_f625w	0.000	75	stromgren_u	1.432
16	**wfpc2_f555w**	0.000	46	acswfc_f775w	0.000	76	stromgren_v	0.179
17	wfpc2_f569w	0.000	47	acswfc_f814w	0.000	77	stromgren_b	0.018
18	wfpc2_f606w	0.000	48	acswfc_f850lp	0.000	78	**stromgren_y**	0.014
19	wfpc2_f656n	0.000	49	acshrc_f220w	0.000	79	tycho_b	0.069
20	wfpc2_f673n	0.000	50	acshrc_f250w	0.000	80	**tycho_v**	0.032
21	wfpc2_f675w	0.000	51	acshrc_f330w	0.000	81	sdss_u	0.042
22	wfpc2_f702w	0.000	52	acshrc_f435w	0.000	82	**sdss_g**	−0.002
23	wfpc2_f785lp	0.000	53	acshrc_f475w	0.000	83	sdss_r	−0.003
24	wfpc2_f791w	0.000	54	acshrc_f550m	0.000	84	sdss_i	−0.016
25	wfpc2_f814w	0.000	55	**acshrc_f555w**	0.000	85	sdss_z	−0.028
26	wfpc2_f850lp	0.000	56	acshrc_f606w	0.000	86	**2mass_j**	−0.021
27	**stisccd_50ccd**	0.000	57	acshrc_f625w	0.000	87	2mass_h	0.009
28	stisccd_f28x50lp	0.000	58	acshrc_f775w	0.000	88	2mass_ks	0.000
29	**stisnuv_25mama**	0.000	59	acshrc_f814w	0.000			
30	stisnuv_f25srf2	0.000	60	acshrc_f850lp	0.000			