`Ulysses`: Principles and practice

prepared by: Tri L. Astraatmadja
affiliation : Max Planck Institute for Astronomy
reference: GAIA-C8-TN-MPIA-TLA-001-1
issue: 1
revision: 7
date: 2015-01-23
status: Issued

Abstract

This Technical Note documents the working principles of `Ulysses`, a new BP/RP simulator written specifically to meet the needs of CU8 to simulate spectra of objects as seen by *Gaia*.

Document History

Issue	Revision	Date	Author	Comment
D	7	2015-01-23	TLA	Adding discussions in Section 1 on XpSim and in Section 3.2 on the polynomial and averaged dispersion curve. Adding an Acknowledgements section.
D	6	2014-10-08	TLA	Seventh draft, rewriting and updating Section 8 after corrections are made to the source code of GOG 14.0.0. Adding Section 8.4 on the comparisons of Ulysses and GOG noisy spectra. Implementing corrections and suggestions from CBJ.
D	5	2014-09-16	TLA	Sixth draft, adding Section 8.2 on the comparison between Ulysses and GOG spectra. Adding an outline of the basic differences between Ulysses and GOG in Section 1.
D	4	2014-07-24	TLA	Fifth draft, adding Appendix B on unit conversion, adding a description of the Fitzpatrick (1999) extinction curve, replacing many example spectra in Appendix C.
D	3	2014-06-30	TLA	Fourth draft, adding Section 7–8 on the noise model and comparisons with GOG.
D	2	2014-03-27	TLA	Third draft, adding Section 6 on noise-free continuous spectra.
D	1	2014-03-18	TLA	Second draft, adding Section 5 on SED normalization, revamping some terms and notations after suggestions by CBJ.
D	0	2014-03-01	TLA	First draft.

Contents

1 Introduction

The need for a BP/RP simulator that can specifically meet the needs of CU8 has long been discussed. CU8 requires a faster turn-around time in the generation of new simulation data, and an ability for the simulation to update the instrument model based on new informations on the satellite (Bailer-Jones, CBJ-066). Ulysses has been developed to meet these needs. It is a simple BP/RP simulator that takes input spectra and produces the observed BP/RP spectra as the output. It is designed to be simple and quick to use, yet to offer flexibility to control the instrument model parameters, apply extinction to input spectra, and to generate noisy data according to *Gaia*'s observing characteristics.

Ulysses is not a new GOG. It is not designed nor intended as a replacement for GOG, for example it is not a simulator to generate objects and simulate all the parameters observable by *Gaia*. The design considerations of Ulysses were as follows:

- Ulysses does not generate sources by itself, but rather takes input spectra provided by the user in order to produce BP/RP spectra.
- Ulysses produces combined sampled BP/RP spectra for a controllable number of transits. It does not simulate astrometric or RVS measurements. It also does not specifically simulate G-band photometry, which is used as an input, but it does provide estimates of $G_{\rm BP}$ and $G_{\rm RP}$ magnitudes derived directly from its BP/RP simulations.
- Ulysses convolves the input spectra with an averaged LSF. GOG users, on the other hand, have to specify the LSF from which row (and for the AF case, the strip) of the CCD and from which telescope is going to be used. While this is appropriate to simulate epoch observations (i.e. single observations), it is less convenient if we want to simulate a combination of multi-epoch observations that span the whole duration of the mission. For this reason Ulysses adopts an averaged instrument model.

The aims of Ulysses to provide BP/RP spectra with faster turnaround time and by simplifying the simulation is in the same spirit as XpSim, a BP/RP simulator written for the development and testing of CU5's PhotPipe algorithms and code (Brown et al., AB-029). However, Ulysses has different designs with XpSim. Ulysses does not simulate spectra image using 2D PSF nor does it simulate the telemetry and CTI effects due to radiation damage. XpSim does not have a noisy model, but as mentioned previously Ulysses does provide noisy spectra.

In Ulysses, the BP/RP prism system is simulated from first principles. The line spread function (LSF) library for all CCDs and both telescopes is taken from the analytical LSF library which are based on instrument simulations. Spectral sampling is performed by numerically integrating all the light that falls on a particular pixel. The final output comprises, among other things,

the wavelength of the sampled spectrum, the noise-free spectrum, and optionally noise-added spectra.

This document describes the principles behind Ulysses. A comparison of Ulysses spectra with GOG spectra is also given.

2 One-dimensional spectra formation in *Gaia*

In this section the mathematical modelling of spectra formation in *Gaia* will be described. Suppose a beam of light from a source with a photon flux SED is $N(\lambda)$ given in photons s^{-1} m^{-2} nm^{-1}. This is the SED before it enters the optical system of *Gaia*, and it already includes the desired absorption effects by interstellar matters and normalized to a given G magnitude. The light passes through the *Gaia* optical system through one of the two telescopes pointing to two different direction separated by an angle of 105°. Each of these two telescopes comprises 6 mirrors, three of which are common to the two telescope. Light from any source is reflected off these 6 mirrors and passes through a filter and a prism before reaching the CCD detector. These components have their own transmissivity as a function of wavelength λ, and the prisms disperse the light into the focal plane of the optical system, which will be recorded by the CCD array. The CCDs also have their quantum efficiencies as a function of λ. These properties are depicted in Figure 1.

At the focal plane, the light is dispersed by the prism along the scanning direction (the AL direction). If we denote the continuous one-dimensional coordinates along this dispersion direction in the data space as κ, then the location in the data space of a reference wavelength λ_0 will be κ_0. The position κ_λ of the monochromatic light with a given wavelength λ can then be given as the offset from the reference point κ_0:

$$\kappa_\lambda - \kappa_0 = \kappa_p(\lambda, \kappa_0), \tag{1}$$

where κ_p is the dispersion curve of the prism considered. The one-dimensional dispersed spectrum $\mathcal{S}$ is then (Brown, AB-009)

$$\mathcal{S}(\kappa) = (D \times H)\tau \int_\lambda N(\lambda) T_m(\lambda) T_p(\lambda) Q(\lambda) L_\lambda(\kappa - \kappa_\lambda) d\lambda, \tag{2}$$

where the quantities in the equation are listed in Table 1. Note that this integration is performed entirely in the data space and thus represents a continuous version of the observed spectra before it is sampled by the detector pixel grid. The actual data will be the discretized version S of $\mathcal{S}$ (Brown, AB-009), which is limited to a window of K pixels in the AL direction:

$$S(k) = \int \mathcal{S}(\kappa)\delta(\kappa - \kappa_0)d\kappa + \epsilon_S(k) \quad k = 0, \ldots, K-1, \tag{3}$$

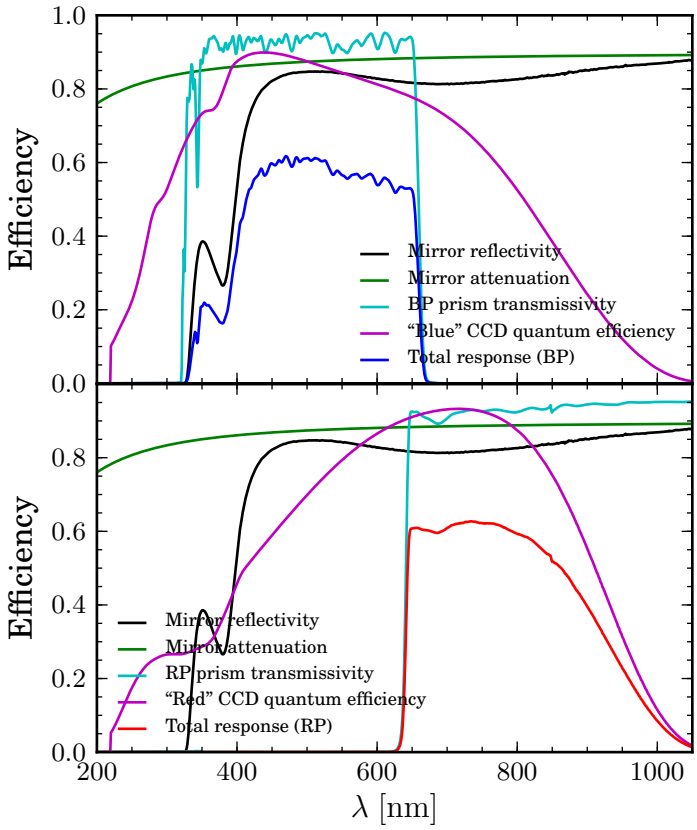


Figure 1: The efficiencies of the *Gaia* optical system in transmitting light through its individual components, as a function of wavelength. The mirror reflectivity (black lines) in this Figure is already raised to the power of the number of mirrors, which is six. The reflectivity attenuation (green lines) of the mirrors is due to contaminations by particulate and molecules as well as the rugosity of the mirrors. The photometer transmissivity (cyan lines) lines in both figures are the throughput of the BP photometer in the top figure, while for the bottom figure it is for the RP photometer. This transmissivity takes into account the throughput of the prism and the filter-coating, as well as reflections losses due to the coatings on the prism. The magenta lines which refer to the quantum efficiency of the CCDs are for the 'blue-enhanced' CCDs for the top figure while for the bottom figure it is for the 'red-enhanced' CCDs.

where $\delta(x)$ is the smearing function in the form of the Dirac delta-function, which has the property

$$\int f(x)\delta(x-a)dx = f(a), \tag{4}$$

and $\epsilon_S(k)$ is the noise added to the data due to the measurement process. This quantity depends on the adopted noise model and will be discussed in more details in Section 7. Note that the smearing function can be in different form. The smearing function adopted in Ulysses is

discussed in Section 6

In the following Section, the quantities in Equation 2 will be described in more details.

3 Modelling the *Gaia* optical system

The optical system of *Gaia* consists of two telescopes, six mirrors, and for the BP/RP spectrographs: one prism each. In the following Section the mathematical modelling of each optical component will be described. Also described in this Section is the modelling of the LSF.

3.1 The telescope and mirrors

A single mirror in *Gaia* has $\sim 97\%$ reflectivity at $\lambda \gtrsim 430$ nm. This $R_{\rm Ag}(\lambda)$ curve can be retrieved from the GPDB under the `Satellite.TELESCOPE_MIRRORREFLECTIVITY` heading. However, as there are 6 mirrors to reflect the detected light toward the focal array, this reflectivity must be raised to the power of the number of mirror. Hence, the total reflectivity of the mirrors $R_m(\lambda) = R^6_{\rm Ag}(\lambda)$. This total reflectivity of the telescopes and the mirrors are shown as the black lines in Figure 1.

The optics will also suffered attenuation due to particulate and molecular contamination and the rugosity of the mirrors. This particulate and molecular contamination will typically reduce the telescope transmissivity by $\sim$ 8–10%. The rugosity of the mirrors is ~ 6 nm. The attenuation due to this combined effects is (EADS Astrium, GAIA.ASF.RP.SAT.00005)

$$R_a(\lambda) = (1 - a_c) \exp\left(-4.6 \left[\frac{2\pi a_r}{\lambda}\right]^2\right), \tag{5}$$

where a_c is the attenuation due to contamination in fractions between 0 and 1, and a_r is the effective rugosity. Both quantities are retrievable from the GPDB under the `Satellite.xP.OPTICS_ATTENUATION_PARTICULATEANDMOLECULARCONTAMINATION` and `Satellite.xP.OPTICS_RUGOSITY_EFFECTIVE` variable name.

The total transmissivity of the telescope and mirrors is then $T_m(\lambda) = R^6_{\rm Ag}(\lambda) R_a(\lambda)$. The GPDB variable names used to derive this transmissivity function is shown in Table 2.

3.2 The BP/RP prism

The modelling of the BP/RP prism is used to calculate the position $x(\lambda)$ of a beam of monochromatic light with wavelength λ on the focal plane, and vice versa. The position $x(\lambda)$ given λ is (Brown, AB-005)

$$x(\lambda) - x_0 = k\left[\sin\delta(\lambda) - \sin\delta_0\right], \tag{6}$$

Table 1: Parameters used in Equation 2 and the GPDB parameters used, if it is available.

$D \times H$	Telescope pupil area (AL $\times$ AC size)	`Satellite.TELESCOPE_PUPILAREA_AL` `Satellite.TELESCOPE_PUPILAREA_AC`
τ	TDI integration time per CCD	`Satellite.CCD_EXPOSURETIME`
$N(\lambda)$	Source SED in photons s^{-1} m^{-2} nm^{-1}	
$T_m(\lambda)$	The total transmissivity of the mirrors	`Satellite.BP.PHOTOMETER_TRANSMISSIVITY`
$T_p(\lambda)$	The total transmissivity of the prism	`Satellite.RP.PHOTOMETER_TRANSMISSIVITY`
$Q(\lambda)$	CCD quantum efficiency at 160 K	`Satellite.BP.CCD_QE_160K` `Satellite.RP.CCD_QE_160K`
L_λ	Monochromatic LSF at wavelength λ	

Table 2: Parameters used for modelling the telescope and mirror system and the GPDB parameters used, if it is applicable.

$R_{\mathrm{Ag}}(\lambda)$	The reflectivity of a single mirror	`Satellite.TELESCOPE_MIRRORREFLECTIVITY`
$R_m(\lambda)$	The total reflectivity of the mirror system	`Satellite.TELESCOPE_TRANSMISSIVITY`
$R_a(\lambda)$	Reflectivity attenuation of the mirrors	
$T_m(\lambda)$	The total transmissivity of the mirrors	$T_m(\lambda) = R^6_{\mathrm{Ag}}(\lambda) R_a(\lambda)$

Table 3: Parameters used for modelling the BP/RP prism and the GPDB parameters used, if it is applicable.

x	Coordinates on the focal plane	
δ	Deviation angle	
k	Proportionality constant	
l_S	Spectrum length	`Satellite.BP.SPECTRUM_LENGTH` `Satellite.RP.SPECTRUM_LENGTH`
λ_0	Reference wavelength	`Satellite.BP.CCD_POSITION_REFERENCEWAVELENGTHOBJECTSPACE` `Satellite.RP.CCD_POSITION_REFERENCEWAVELENGTHOBJECTSPACE`

$\lambda_{\mathrm{max,min}}$	The minimum and maximum wavelength of the BP/RP spectra	`Satellite.BP.WAVELENGTHRANGE_MINIMUM`
		`Satellite.BP.WAVELENGTHRANGE_MAXIMUM`
		`Satellite.RP.WAVELENGTHRANGE_MINIMUM`
		`Satellite.RP.WAVELENGTHRANGE_MAXIMUM`

α	Top angle of the prism	`Satellite.BP.PRISM_ANGLE` `Satellite.RP.PRISM_ANGLE`
θ	Incidence angle of the beam	
n	Refractive index	

$B_{1,2,3}$	Sellmeier coefficients B	`Nature.FUSEDSILICA_SELLMEIERCOEFFICIENT_B1`
		`Nature.FUSEDSILICA_SELLMEIERCOEFFICIENT_B2`
		`Nature.FUSEDSILICA_SELLMEIERCOEFFICIENT_B3`

$C_{1,2,3}$	Sellmeier coefficients C	`Nature.FUSEDSILICA_SELLMEIERCOEFFICIENT_C1`
		`Nature.FUSEDSILICA_SELLMEIERCOEFFICIENT_C2`
		`Nature.FUSEDSILICA_SELLMEIERCOEFFICIENT_C3`

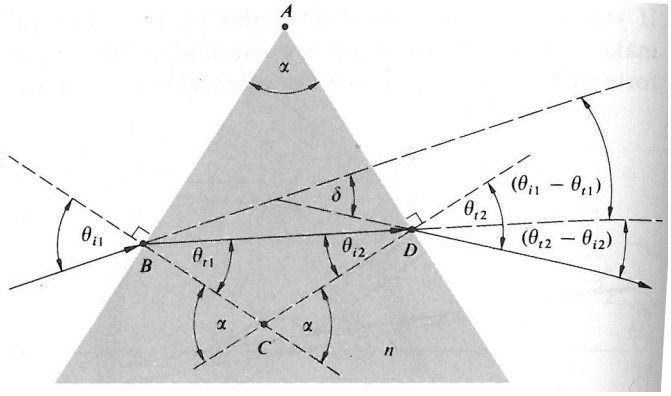


Figure 2: The geometry of a dispersing prism. Here α is the top angle of the prism, θ_{i1} is the incidence angle of a beam of light, and δ is the total deviation angle of the beam. The deviation angle is the angular difference between the original path of the beam and the deflected path of the beam after it exits the prism. Figure reproduced from Hecht (2001).

where $\delta(\lambda)$ is the deviation angle of the light in question, δ_0 is the deviation angle of a reference wavelength λ_0, x_0 is the position of that light with reference wavelength on the focal plane, and k is the proportionality constant which can be calculated by defining the pixel coordinates of the blue and red wavelength edges of the spectra (Brown, AB-005). We can do this by dividing the average spectrum length $\langle l_S \rangle$ with the distance on the focal plane between the maximum wavelength and the minimum wavelength:

$$k = \frac{\langle l_S \rangle}{\sin\delta(\lambda_{\max}) - \sin\delta(\lambda_{\min})}. \tag{7}$$

The deviation angle δ for a beam of light with wavelength λ coming towards the prism at an angle of incidence θ is (Hecht, 2001)

$$\delta(\lambda) = \theta + \sin^{-1}\left[(\sin\alpha)(n^2 - \sin^2\theta)^{1/2} - \sin\theta\cos\alpha\right] - \alpha, \tag{8}$$

where α is the top angle of prism and n is the refractive index of the prism as a function of λ. δ increases with n, which is a function of wavelength. Thus δ is a function of wavelength as well. In Figure 3 we can see that the refractive index n of the prism used in *Gaia* decreases as the wavelength λ increases. The deviation angle $\delta(\lambda)$ will thus be less for red light than it is for the blue light. The deviation angles for both prisms are shown in Figure 4.

The prisms used for the BP/RP spectrograph system are made of fused silica of type SIL120 (Domingues & Rebordao, CDO-001). Its refractive index n as a function of λ can be described by the Sellmeier equation

$$n^2 = 1 + \sum_{i=1}^{3} \frac{B_i\lambda^2}{\lambda^2 - C_i}, \tag{9}$$

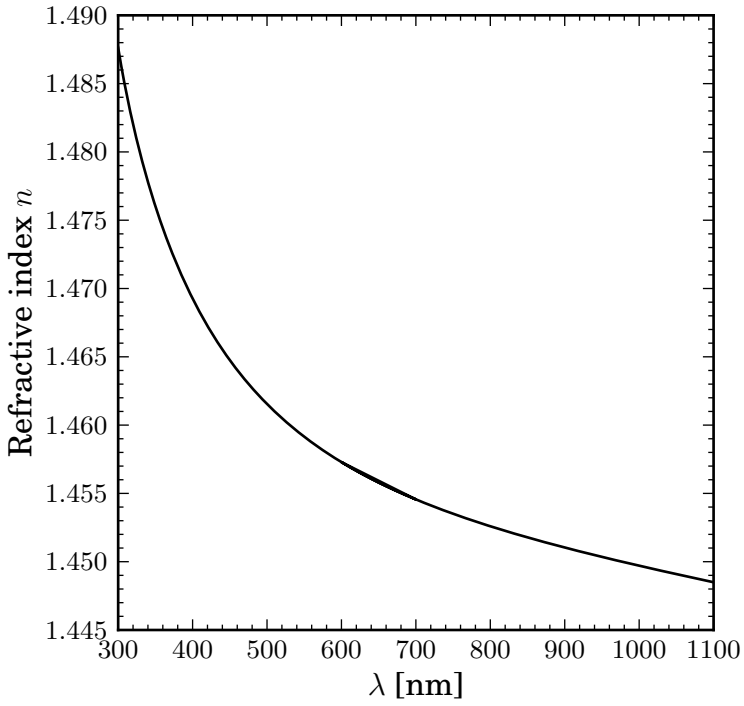


Figure 3: The refractive index of the BP/RP prism as a function of λ. The material of the prism is SIL120 fused-silica.

where B_i and C_i are the Sellmeier Coefficients for the material considered, and are experimentally determined. In the GPDB, the Sellmeier Coefficients are defined for working temperature $T = 120$ K and λ measured in μm. The refractive index of the prism as a function of λ is shown in Figure 3.

If the position on the focal plane $x(\lambda)$ is known, the spectral sampling can then be calculated by (Brown, AB-005)

$$\frac{d\lambda}{dx} = \left(\frac{dx}{d\lambda}\right)^{-1} \text{ nm pixel}^{-1}, \tag{10}$$

where

$$\frac{dx}{d\lambda} = k \cos\delta \frac{d\delta}{d\lambda}, \tag{11}$$

and

$$\frac{d\delta}{d\lambda} = \left(1 - \left[\sin\alpha(n^2 - \sin^2\theta)^{1/2} - \sin\theta\cos\alpha\right]^2\right)^{-1/2} \frac{n\sin\alpha}{(n^2 - \sin^2\theta)^{1/2}} \frac{dn}{d\lambda}, \tag{12}$$

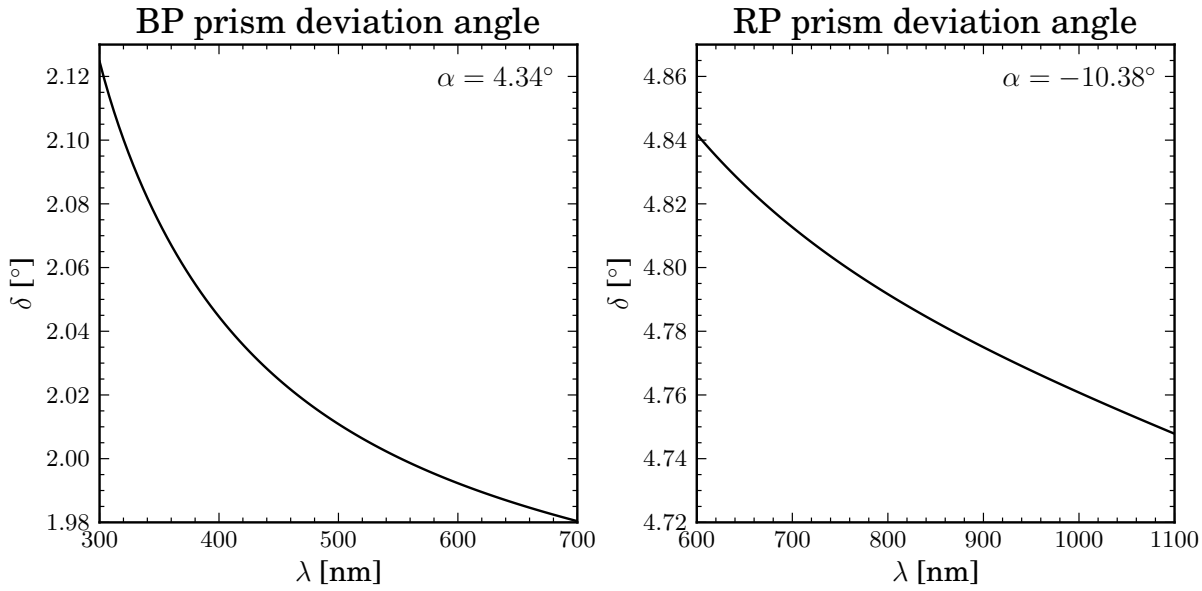


Figure 4: The deviation angle δ of the BP (left plot) and RP (right plot) prism, as a function of wavelength λ. The different shape of the $\delta(\lambda)$ curves are due not only to the variation the refractive index n with regard to λ at different regimes of wavelength, but also because of the different top angle α of the prisms: The BP prism has $\alpha = 4.34°$, while the RP prism has $\alpha = -10.38°$. Here the minus sign of the top angle of the RP prism indicates that the dispersion direction is to the opposite direction.

where

$$\frac{dn}{d\lambda} = \frac{1}{2n}\frac{d}{d\lambda}(n^2 - 1) \tag{13}$$

$$= -\frac{1}{2n}\sum_{i=1}^{3}\frac{2B_i C_i \lambda}{\lambda^2 - C_i}. \tag{14}$$

Thus using Equation 6 and the subsequent equations, we can calculate the dispersion curve, i.e. the position of the monochromatic light with wavelength λ on the focal plane. The dispersion curves for the BP/RP prism, along with the spectral sampling, are shown in Figure 5. In these calculations, the incidence angle θ is assumed to be zero.

We can also work out the inverse problem, i.e. to calculate the corresponding wavelength λ, given the position of that monochromatic light x_λ on the focal plane. This is useful to work out the original wavelength that form the spectrum after we perform the calculations in Equation 2 and 3. Moreover, it is also useful to ensure that our calculations in Equations 6–9 are correct. The wavelength λ of a monochromatic light falling on the focal plane at given position x_λ can

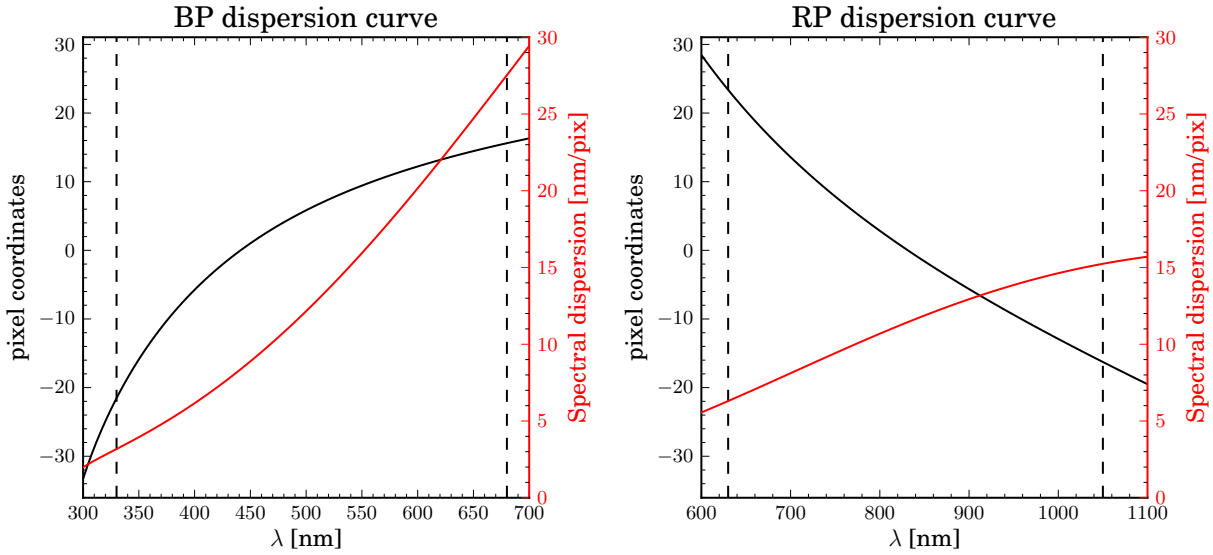


Figure 5: The dispersion curve (black lines) and spectral sampling (red lines) of the BP/RP, calculated using Equations 6 and 10. The vertical dashed-lines indicate the blue and red wavelength limits of the spectra.

be calculated first by calculating the deviation angle, given x_λ:

$$\delta(x_\lambda) = \sin^{-1}\left(\frac{x_\lambda - x_0}{k} + \sin\delta_0\right). \tag{15}$$

Given the deviation angle δ, we can then calculate the refractive index of the prism, if we also know the incidence angle θ and the prism top angle α:

$$n = \sqrt{\sin^2\theta + \left[\frac{\sin\theta\cos\alpha + \sin(\delta - \theta + \alpha)}{\sin\alpha}\right]^2}. \tag{16}$$

We can calculate λ from the refractive index n by inverting the Sellmeier equation. We can do this using numerical methods such as bisection or Newton-Raphson, but the quickest way (in terms of computational time) is by working out the inverse function of the Sellmeier equation. We do this first by rearranging the equation into a cubic function:

$$\begin{aligned} 0 = {} & [B_1 + B_2 + B_3 - (n^2 - 1)]\lambda^6 \\ & + [(C_1 + C_2 + C_3)(n^2 - 1) - B_1(C_2 + C_3) - B_2(C_1 + C_3) - B_3(C_1 + C_2)]\lambda^4 \\ & + [B_1C_2C_3 + B_2C_1C_3 + B_3C_1C_2 - (C_1C_3 + C_2C_3 + C_1C_2)(n^2 - 1)]\lambda^2 \\ & + C_1C_2C_3(n^2 - 1), \end{aligned} \tag{17}$$

which could then be simplified as

$$z = \lambda^2, \tag{18}$$
$$K = [B_1 + B_2 + B_3 - (n^2 - 1)], \tag{19}$$
$$a_0 = \frac{1}{K}[C_1 C_2 C_3 (n^2 - 1)], \tag{20}$$
$$a_1 = \frac{1}{K}[B_1 C_2 C_3 + B_2 C_1 C_3 + B_3 C_1 C_2 - (C_1 C_3 + C_2 C_3 + C_1 C_2)(n^2 - 1)], \tag{21}$$
$$a_2 = \frac{1}{K}[(C_1 + C_2 + C_3)(n^2 - 1) - B_1(C_2 + C_3) - B_2(C_1 + C_3) - B_3(C_1 + C_2)]. \tag{22}$$

Applying these definitions, Equation 17 then becomes

$$z^3 + a_2 z^2 + a_1 z + a_0 = 0. \tag{23}$$

This is obviously a cubic equation which can be solved for z through Vieta's substitution[1]. This is done first by defining intermediate variables

$$Q \equiv \frac{1}{9}(3a_1 - a_2^2), \tag{24}$$
$$R \equiv \frac{1}{54}(9a_2 a_1 - 27a_0 - 2a_2^3), \tag{25}$$
$$D \equiv Q^3 + R^2. \tag{26}$$

Here D is the polynomial discriminant which tell us the properties of the roots of Equation 23. If $D > 0$, one root is real while the other two are complex conjugates. If $D = 0$, all roots are real and at least two are equal, and if $D < 0$, all roots are real and unequal. Due to the nature of the Sellmeier equation and its coefficients, D will always be greater than zero. Hence there is one real root to Equation 17, and this can be solved by first defining

$$S \equiv \sqrt[3]{R + \sqrt{D}}, \tag{27}$$
$$T \equiv \sqrt[3]{R - \sqrt{D}}. \tag{28}$$

The real root of z is then

$$z = -\frac{1}{3}a_2 + S + T. \tag{29}$$

The wavelength λ can then be found if we take the square root of z:

$$\lambda = 1000\sqrt{z(n)}\ \mathrm{nm}. \tag{30}$$

Hence these two sets of equations, to calculate $x(\lambda)$ given λ and vice versa, can be used to produce two dispersion curves (Figure 5) which should have exacly the same shapes, provided

[1] http://mathworld.wolfram.com/CubicFormula.html

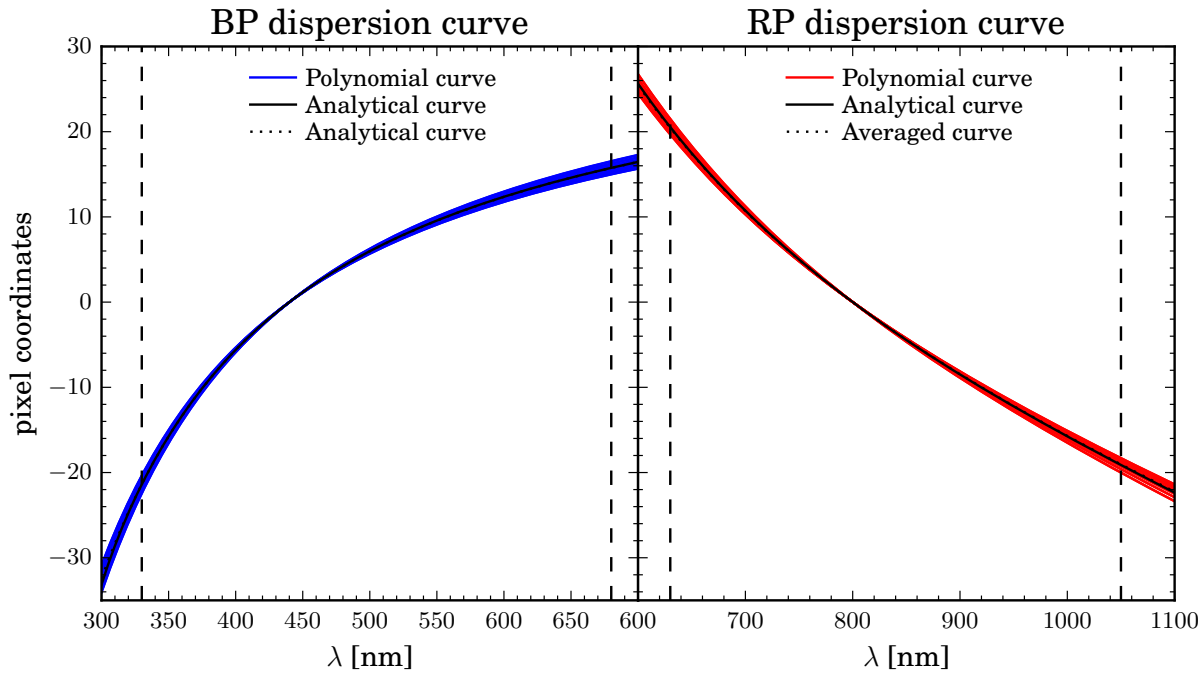


Figure 6: The dispersion curves of the BP/RP prism. The blue and red lines correspond to the polynomial dispersion curve calculated from the fitting of the Zemax data. The black solid-lines are the dispersion curves calculated from first principles, and the green dashed-lines are the averaged dispersion curves calculated by averaging all the polynomial curves. The green dashed-lines are mostly on top of the black solid-lines, which indicates that the analytical dispersion curve is a good approximation of an averaged dispersion curve.

the coding is correct. The calculation of λ given pixel coordinates x is also useful to transform the pixel coordinates into wavelengths.

Having worked through this analytical equations of the prism model, one can ask how these analytical model shown in Figure 5 differs with the dispersion curve derived from the fitting of the ray-tracing data generated with Zemax. In Domingues & Rebordao (CDO-001), the polynomial coefficients for the dispersion curve is calculated for a univariate model (the dispersion depends only on wavelength) and a bivariate model (the dispersion depends on the field angle in the AC direction as well). The polynomial coefficients of the dispersion curve are also calculated for individual BP/RP chips and FOV. Hence, for BP there are a total of 14 dispersion curves, one for each chip and one for each FOV. The same also goes for RP. These coefficients are retrievable from the GPDB.

I took the coefficients for the univariate dispersion curve from the GPDB and calculate the curve for each chip and FOV. The proportionality constant k for each dispersion curve is still calculated using Equation 7, but instead of using an averaged spectrum length the spectrum length for each

chip used.

The plot of all dispersion curves for BP and RP is shown in Figure 6. As we can see, the dispersion curves for each chip and FOV covers a certain "band" (shown as the red and blue lines that are so close to each other that they resemble a band) which shows the possible pixel coordinates given the wavelength, row number, and the telescope number. If we average these dispersion curves to obtain an averaged curve, the curve will be almost on top of the analytical dispersion curve calculated from first principles (the averaged dispersion curve is shown as the green dashed-lines in Figure 6). Thus we can be sure that the analytical dispersion curve is an excellent approximation to the averaged dispersion curve.

Light also suffer attenuation when passing through the prisms. The transmissivity curves of both prisms are shown in Figure 1 as the cyan lines. Both transmissivity curves already take into account the low-pass filter coatings.

3.3 The line spread function (LSF)

The distribution of the monochromatic light of wavelength λ has on the focal plane can be described by a two-dimensional function called the point-spread function (PSF), or by its collapsed function to one dimension called the line-spread function (LSF). Both functions have been generated using instrument simulations (Gardiol et al., DG-011). Since we are interested in simulating only one-dimensional spectra, using the LSF is enough for this purpose.

The *Gaia* LSF have been simulated for the SM, AF, BP, RP, and RVS focal plane. For BP and RP, there is 1 LSF for each CCD (7 CCD rows for each prism) and telescopes (2 telescopes) combination. The LSF is sampled in λ at 25 nm interval from 325–700 nm for BP and 625–1050 nm for RP, and at 3 points per pixel (Busonero & Gardiol, DB-007). Each LSF is then fitted with the bi-quartic B-spline function to represent the core of the LSF, and the Cauchy distribution to represent the wings (Lindegren, GAIA-LL-046). The bi-quartic B-spline function has 31 knots, while the Cauchy function has 3 coefficients. The values of these coefficients for all LSF representing the CCD chips can retrieved at `http://gaia.am.ub.es/CU2/Data-13.0/lsf/`.

The plot of all these LSF are shown in Figures 7 and 8. All of the coefficients will produce unnormalized LSFs, so it is necessary to normalize these LSFs such that $\int_{-\infty}^{+\infty} L_\lambda(x - x(\lambda))dx = 1$, before using them further.

To simplify the simulations, we choose to adopt, for each of BP and RP, a single LSF for both telescopes and all CCDs. This is done by averaging all the relevant LSFs. These averaged LSF for the nominal wavelengths are then used to interpolate the LSF at any other wavelengths. The bicubic-spline interpolation is used for this. The results of these averaging and interpolation are shown in Figures 9 and 10. Once we are able to obtain the LSF for all wavelengths, we can then

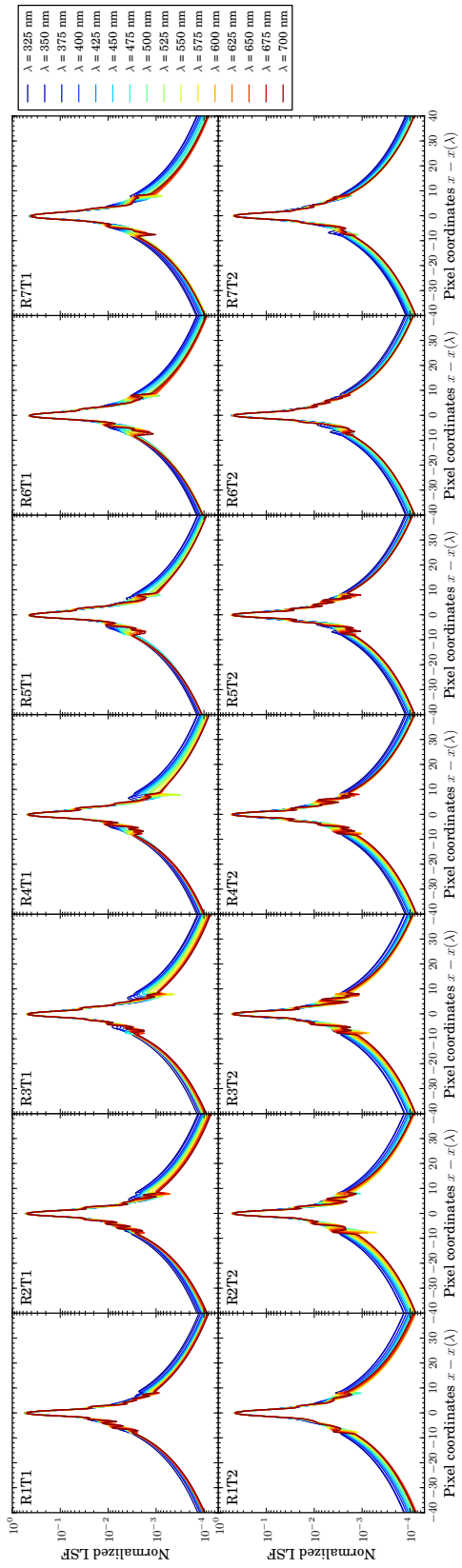


Figure 7: The normalized LSF for all rows of the BP CCDs and for light coming from both telescopes, as a function of offset in pixel coordinates relative to the position of the light at wavelength λ. For each plot, the LSF is plotted for the wavelengths color-coded according to the legend.

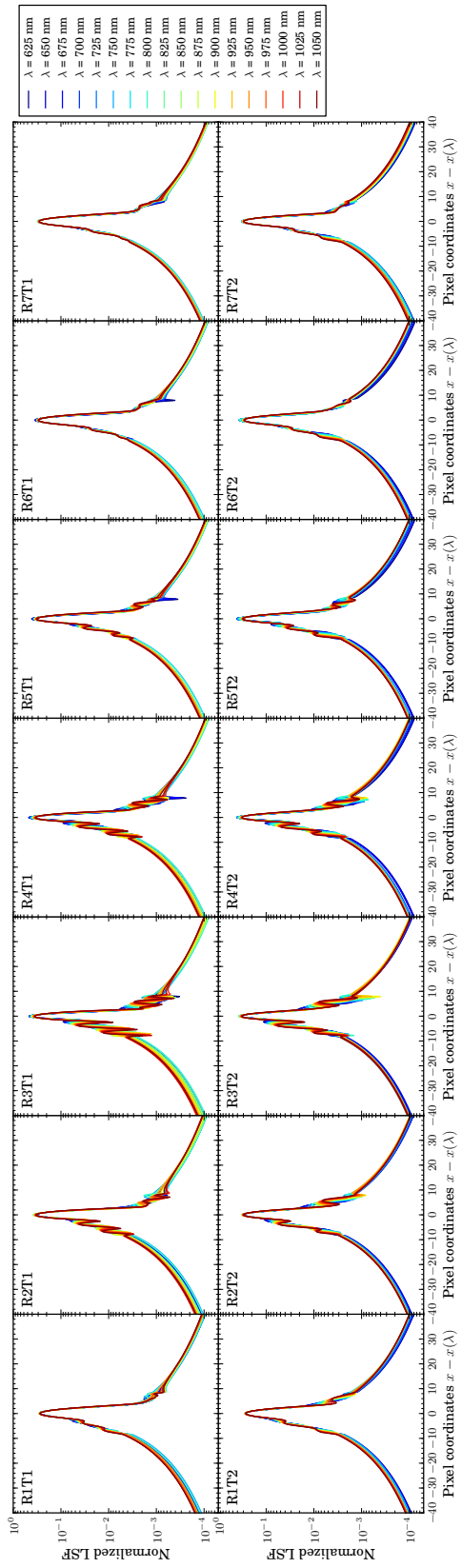


Figure 8: The normalized LSF for all rows of the RP CCDs and for light coming from both telescopes, as a function of offset in pixel coordinates relative to the position of the light at wavelength λ. For each plot, the LSF is plotted for the wavelengths color-coded according to the legend.

integrate Equation 2.

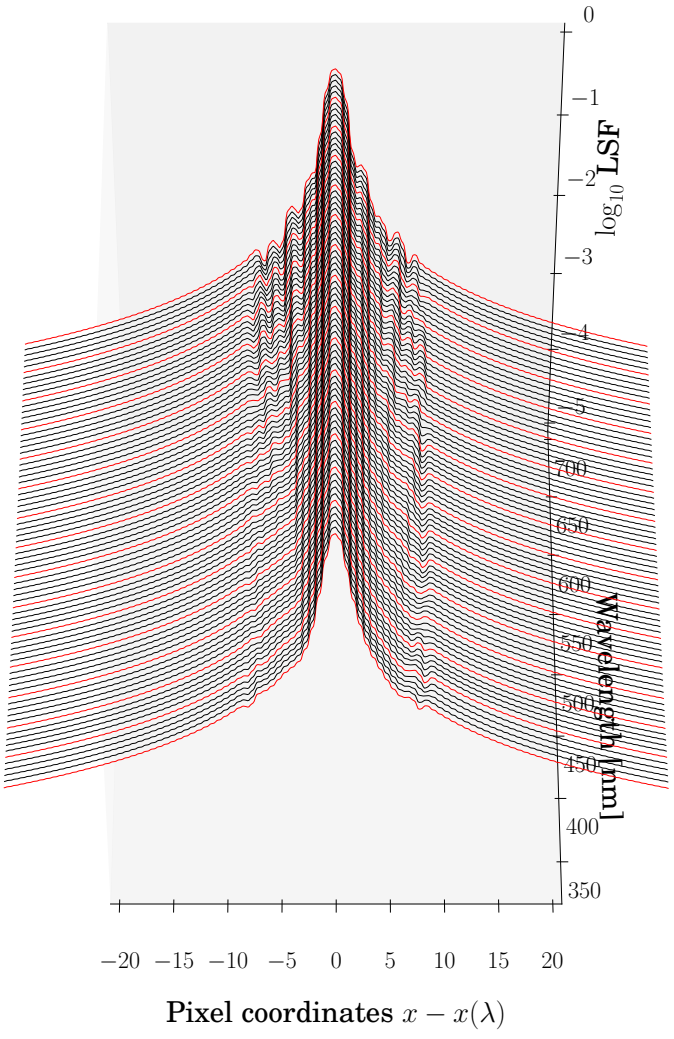


Figure 9: The BP normalized mean LSF. Red lines are for the LSF at nominal wavelength, while the black lines are the interpolated LSF.

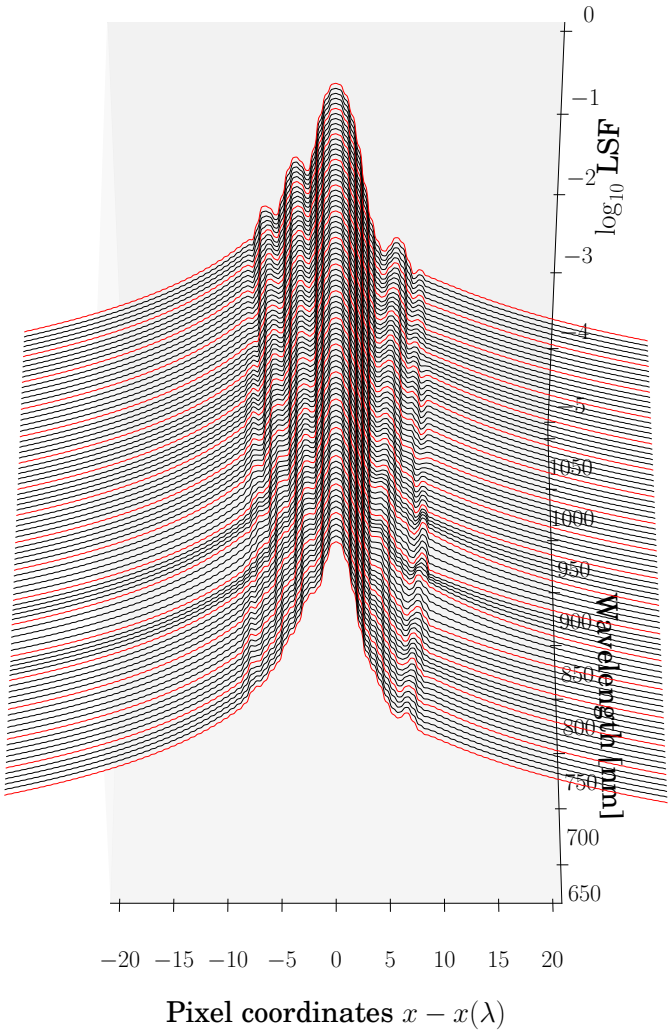


Figure 10: The RP normalized mean LSF. Red lines are for the LSF at nominal wavelength, while the black lines are the interpolated LSF.

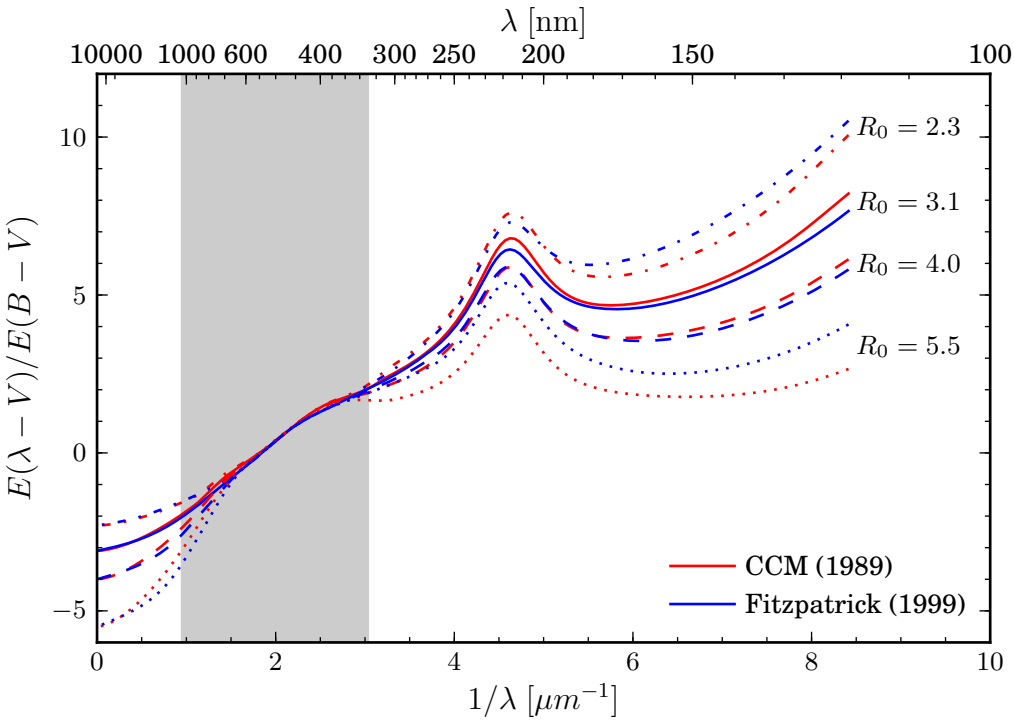


Figure 11: A comparison between the Cardelli et al. (1989, CCM) and the Fitzpatrick (1999) normalized extinction curve for different values of R_0, from the far-infrared though the UV. The normalization is such that $E(\lambda - V)/E(B - V) = -R_0$ as $1/\lambda \to 0$. The shaded part of the plot is the *Gaia* sensitivity range.

We can see in Equation 2 and from Figures 7–10 that the LSFs spread the light with wavelength λ, that is supposed to fall into position $\kappa(\lambda)$, into the neighboring pixels as well. Thus at a particular AL position $\kappa(\lambda)$, the light from all wavelengths in the *Gaia* wavelength range actually contributes to the total flux at that position, although the majority of the contribution will come from the light with wavelength λ (Brown, AB-005).

4 Extinction curve

The attenuation of light by interstellar matter can be applied to the input spectra by providing the desired value of A_0, which is the amount of extinction at a monochromatic wavelength of $\lambda_{\rm ref} = 549.45$ nm. While in the past A_0 has been commonly written as A_V, this is actually quite confusing since A_0 is not the extinction in the V-band (Bailer-Jones, 2011). The extinction in the V-band, or any filter for that matter, with a passband function h_λ is defined as the integration of the extinction-affected stellar SED N_λ (the energy flux of a source as a function of λ) over the

wavelength range of the band, and thus depends on the intrinsic parameters of the star:

$$A_V = -2.5 \log_{10} \left(\frac{\int_V N_\lambda h_\lambda 10^{-0.4 A_\lambda} d\lambda}{\int_V N_\lambda h_\lambda d\lambda} \right). \tag{31}$$

For the same value of A_0, two stars with different intrinsic characteristics can have different values of A_V. In contrast, A_0 depends solely on the properties of the interstellar medium.

Ulysses employs the commonly-used extinction curve models of Cardelli et al. (1989) and Fitzpatrick (1999). Users can select which model to use in the Ulysses property file before running a simulation. Both extinction curves for various values of R_0 are shown in Figure 11.

From the extinction curve in Figure 11 we can determine the extinction at any given wavelength $A(\lambda)$ relative to A_0. Such an extinction curve, normalised to $A_0 = 1$, is shown in the top plot of Figure 12. How such extinction curve would affect any given SED is shown in the bottom plot of the same Figure, in which the attenuation factor as a function of wavelength λ is shown. A comparison between the Cardelli et al. (1989) and Fitzpatrick (1999) extinction curve is also shown.

5 Normalizing the spectral energy distribution

Suppose we already obtain, e.g. from a spectral library, a spectral energy distribution of a source. This SED is already converted into units of photons m^{-2} s^{-1} nm^{-1}. We then apply to this spectrum the effects of interstellar matter absorption, by adopting first a model of an extinction curve as described in Section 4.

The extinction curve in Figure 12 plots the ratio A_λ/A_0, i.e. the absorption at any given wavelength λ, A_λ, relative to the absorption at the visual wavelength A_0. Since the curve is already normalized such that $A_0 = 1$, we can simply write the ratio A_λ/A_0 as $A(\lambda)$. Using this curve, we can then calculate a reddened SED $N(\lambda)$ for any given A_0, given an unreddened SED $N_0(\lambda)$:

$$N(\lambda) = 10^{-0.4 A_0 A(\lambda)} N_0(\lambda). \tag{32}$$

After these effects of interstellar absoption have been applied, we can then normalize the spectrum to any given G magnitude that we want. The G-band magnitude is defined as the integration of the unfiltered white light of a detected source on the AF chips (Jordi et al., 2006, 2010). The G magnitude passband is thus defined by the quantum efficiency of the AF chip and the transmissivity of the mirrors. The zero point of the G-band is calculated by assuming that the magnitude of a Vega-like star is $G = 0.03$ (Jordi et al., 2010). The passband of the G-band, along with the G_{BP} and the G_{RP}-band, is shown in Figure 13, while their total responses are shown in Figure 14.

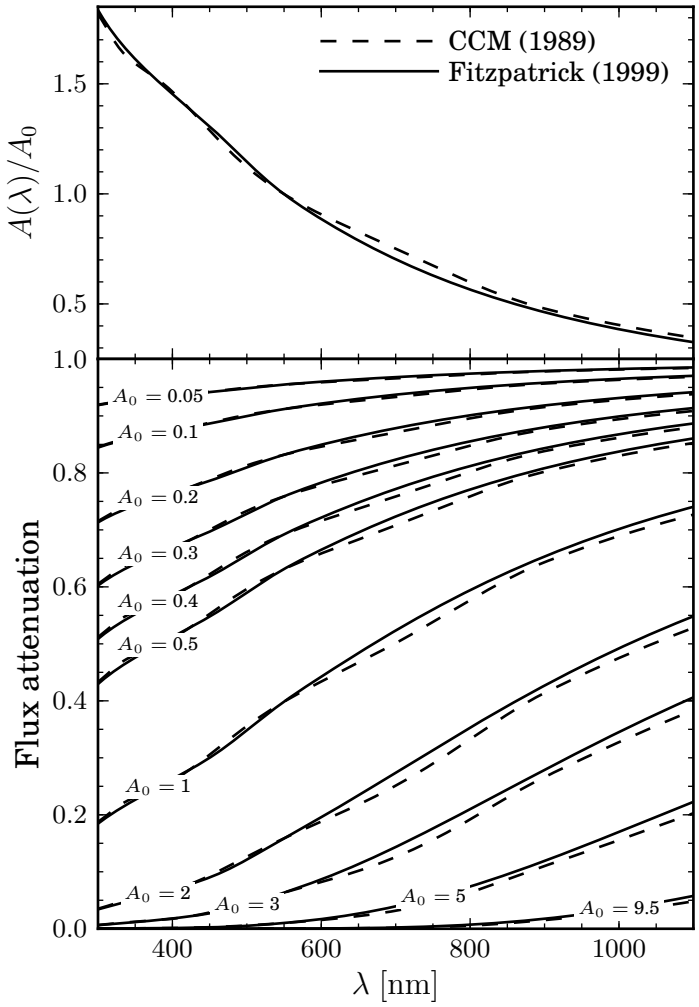


Figure 12: The top plot displays the extinction curve of both Cardelli et al. (1989) and Fitzpatrick (1999) at wavelength range $300 \text{ nm} \leq \lambda \leq 1100 \text{ nm}$, while the bottom plot displays the attenuation factor suffered by any flux at the same wavelength range. Here A_0 is the monochromatic absorption at $\lambda_{\text{ref}} = 550 \text{ nm}$, and the extinction curve at the top plot is normalized to $A_0 = 1$. For both curves, $R_0 = 3.1$

The zero point of each band has been calculated and is available in the GPDB. The normalization of an SED $N(\lambda)$ to any given G-magnitude is then

$$G = -2.5 \log c_N s + G_0, \tag{33}$$

where G_0 is the zero point of the G-band, c_N is the normalization constant of the SED, and

$$s = (D \times H) \int N(\lambda) T_m(\lambda) Q_{\text{AF}}(\lambda) d\lambda, \tag{34}$$

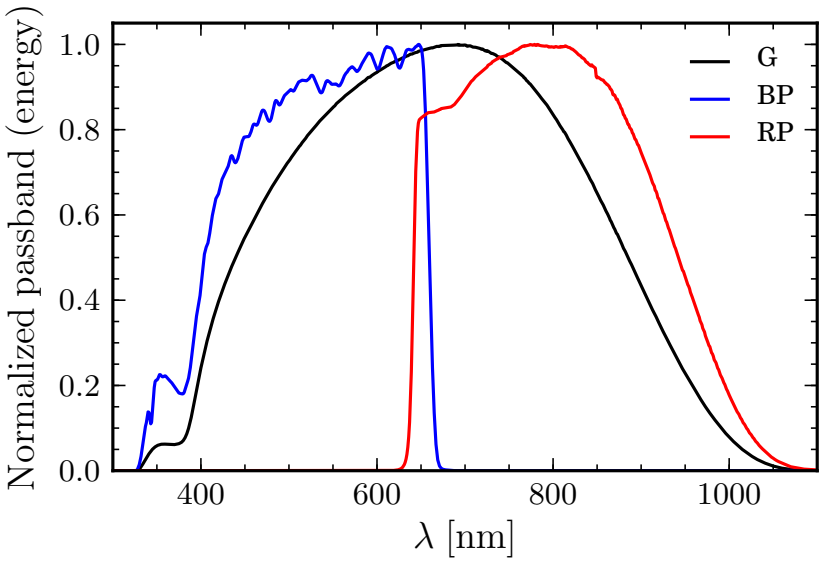


Figure 13: The normalized passbands of the *Gaia* G (black line), $G_{\rm BP}$ (blue line), and $G_{\rm RP}$ (red line) magnitudes. These passbands are calculated using values obtained from the GPDB. The passbands $S_X(\lambda)$ are defined by $S_X(\lambda) = \lambda T_m(\lambda) P_X(\lambda) Q_X(\lambda)$, where $T_m(\lambda)$ is the telescope transmissivity, $P_X(\lambda)$ is the prism transmissivity ($P_X(\lambda) = 1$ for the G-band), and $Q_X(\lambda)$ is the CCD quantum efficiency. The subscript X stands for the G, $G_{\rm BP}$, or the $G_{\rm RP}$ band.

is the integration of the white light from the source which includes the optical system and detector effects, in units of e^- s^{-1}. The normalization constant c_N is thus

$$c_N = \frac{1}{s} 10^{-0.4(G-G_0)}. \tag{35}$$

Multiplying c_N with the reddened spectrum $N(\lambda)$, we then obtain a spectrum normalized to a given G magnitude.

6 Continuous noise-free BP/RP spectra and its discrete sampling

We have described all the quantities necessary to simulate a continuous spectrum by means of Equation 2. The result of such calculation using a flat SED (constant number of photons s^{-1} m^{-2} nm^{-1}) as input, normalized to $G = 15$, is shown in Figure 15. The numerical integration in Equation 2 uses equal width quadratures of degree 4, i.e. Boole's rule (Press et al., 1992).

Since we are using a flat SED, The resulting spectra as observed by *Gaia* in Figure 15 show the overall response of the optical–detector chain of the photometers after it was convoluted with the

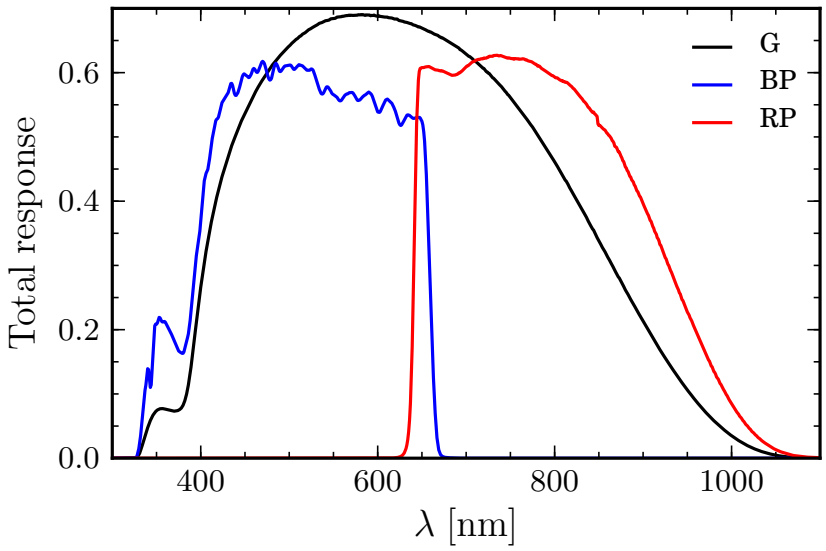


Figure 14: The total response of the *Gaia* G (black line), $G_{\rm BP}$ (blue line), and $G_{\rm RP}$ (red line) bands. These are calculated using the same values as in Figure 13. The response of the band $S_X(\lambda)$ are defined by $S_X(\lambda) = T_m(\lambda)P_X(\lambda)Q_X(\lambda)$. The meaning of these symbols are explained in Figure 13.

detector's LSF. In the left and middle plots of Figure 15, the spectra are shown as a function of the data space coordinates κ in pixel units, while in the right plot both spectra are shown as a function of wavelength. Note that the RP prism disperses light to the opposite direction of the BP prism, as can be seen in the secondary x-axis on the top of the left and right plot in the Figure.

The continuous spectrum that falls on the focal plane array of *Gaia* will be sampled by the discrete array of pixels that comprise the CCD. The sampling method of the pixel will be described in the following, as well as the oversampling method.

The discretization of the continuous spectrum on a pixel grid is performed by integrating the continuous spectrum $\mathcal{S}(\kappa)$ over the interval corresponding to each pixel k (Brown, AB-009):

$$S(k) = \int_{k-0.5}^{k+0.5} \mathcal{S}(\kappa)\Pi(\kappa - k)d\kappa, \tag{36}$$

where $\Pi(x)$ is the pixel smearing function defined as

$$\Pi(x) = \begin{cases} \frac{1}{2} & \text{for } x = \pm 0.5, \\ 1 & \text{for } -0.5 < x < +0.5, \\ 0 & \text{elsewhere.} \end{cases} \tag{37}$$

Here we adopt the FITS convention by placing the centres of pixels at integer values of the

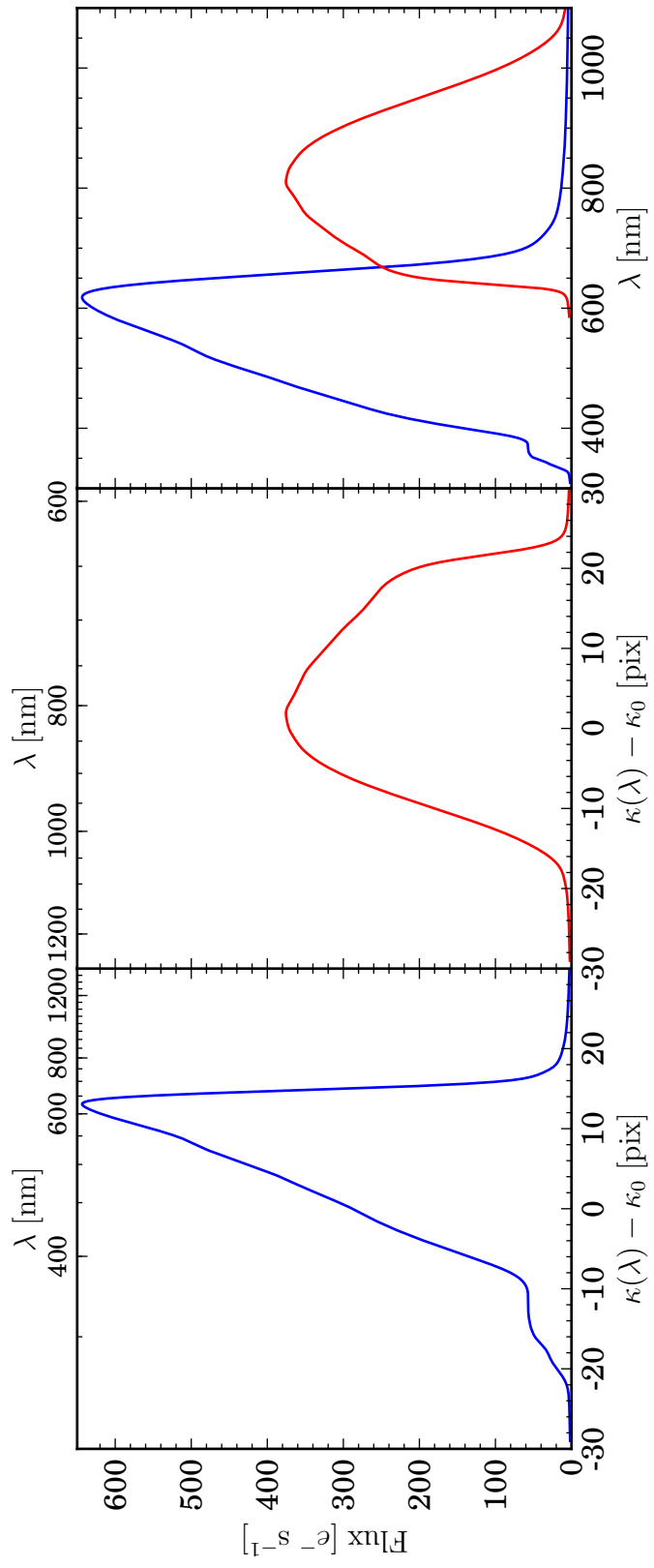


Figure 15: The BP/RP continuous spectra for a source with a flat SED (constant number of photons s^{-1} m^{-2} nm^{-1}) normalized to magnitude $G = 15$. The left and middle plots show respectively the BP and RP spectra as a function of coordinates in data space, where the zero point is centered on the respective reference wavelength. The corresponding wavelength for each spectrograph is displayed on the secondary x-axis at the top of the plot. Note that the dispersion direction of RP is the reverse of BP, as indicated by the wavelength scale on the secondary x-axis. On the right plot, the BP and RP spectra as a function of wavelength is shown. These spectra are noise-free.

continuous coordinates system κ and the size of each pixel is one unit.

When we are oversampling the spectrum by a factor of $n_{\rm over}$, this would mean that we are observing the spectrum at least $n_{\rm over}$ times and we shift the position of the wavelengths on the focal plane in the AL direction by

$$x_i(\lambda) = x_0(\lambda) + \frac{i}{n_{\rm over}}, \qquad i = 0, \ldots, n_{\rm over} - 1. \tag{38}$$

The resulting spectra are then interleaved to obtain a spectrum with a size of the window size multiplied by $n_{\rm over}$.

An example is as the following (Brown, AB-006): Suppose the window size of the spectra is 60 pixels. A normally sampled spectrum ($n_{\rm over} = 1$) is sampled at the exact centre of each pixels which correspond to positions $x = \{0, 1, 2, 3, \ldots, 59\}$. When we oversample the spectrum 4 times in the AL direction this would mean that we sample the spectrum first at the previously mentioned centre of each pixels, then at sample fractions 0.25 at positions $x = \{0.25, 1.25, 2.25, \ldots, 59.25\}$, at sample fractions 0.5 at positions $x = \{0.5, 1.5, 2.5, \ldots, 59.5\}$, and at sample fractions 0.75 at positions $x = \{0.75, 1.75, 2.75, \ldots, 59.75\}$. We then interleave these four spectra to obtain a spectrum containing 60×4 pixels corresponding to positions $x = \{0, 0.25, 0.5, 0.75, 1, 1.25, \ldots, 59, 59.25, 59.5, 59.75\}$. We can extract a normally sampled spectrum from this 4-oversampled spectrum by taking every fourth data point (e.g $x = \{0, 1, 2, \ldots, 59\}$ or $x = \{0.75, 1.75, 2.75, \ldots, 59.75\}$), or a twice-oversampled spectrum by taking every second data point (e.g. $x = \{0, 0.5, 1, 1.5, 2, \ldots, 59, 59.5\}$).

For additional illustrations, the spectra of several stars with different APs are also simulated and shown in Appendix C. The spectra shown in Figures 33–58 are noise-free and are oversampled by a factor of 8.

7 The noise model

The noise model of Ulysses adopts the model written in Jordi et al. (CJ-043). At the moment Ulysses only supports the simulation of the noise-free spectra and the end-of-mission noisy-spectra.

The sampled, noisy, end-of-mission spectra are generated by first generating the sampled noise-free spectra, then we add the noise by

$$f_{\rm noisy}(k) = f_{\rm noise-free}(k) + r_G(0|1)\sigma_{{\rm EOM},k}, \tag{39}$$

in which $r_G(0|1)$ is a random number drawn from a Gaussian distribution of zero mean and 1 unit of standard deviation, and $\sigma_{{\rm EOM},k}$ is the standard deviation of the noise-free flux measurement at pixel index k.

The end-of-mission standard deviation $\sigma_{\mathrm{EOM},k}$ is calculated by

$$\sigma_{\mathrm{EOM},k} = \sqrt{m^2\sigma_{\mathrm{str}}^2 \frac{n_{\mathrm{over}}}{n_{\mathrm{tr}}} + \sigma_{\mathrm{cal}}^2}, \tag{40}$$

where n_{over} is again the number of oversampling, n_{tr} is the number of transit the source have, i.e. the number of times the object is observed, σ_{str} is the standard deviation of the flux from a single observation of the object, σ_{cal} is the uncertainty of the flux internal calibration, and the m parameter is the overall mission margin created to take into account unknown sources of errors. The value $m = 1.2$ is adopted as a general consensus within *Gaia*. In this equation we can see that the more oversampled a spectrum is, the higher the uncertainty in the flux measurement. However, as a trade-off we obtain more data points in the spectrum.

The uncertainty of the single-transit measurement is a combination of the poisson noise and the total CCD noise of the flux measurement and the background measurement in pixel index k:

$$\sigma_{\mathrm{str}}(k) = \frac{1}{\tau}\sqrt{\sigma_f^2(k) + \sigma_{\mathrm{bg}}^2(k)}, \tag{41}$$

where τ is the TDI integration time, $\sigma_f(k)$ is the uncertainty of the flux measurement, and $\sigma_{\mathrm{bg}}^2(k)$ is the uncertainty of the background measurement. The uncertainty in the flux measurement is

$$\sigma_f^2(k) = (f(k) + n_{\mathrm{bg}})\tau + r^2, \tag{42}$$

Here $f(k)$ is the noise-free flux at pixel index k, n_{bg} is the measured background, and r is the total noise of the CCD. The total CCD noise includes the CCD readout noise (RON), CCD dark noise, kTC noise (the thermal noise on CCD capacitors), overall video-chain noise (including ADC noise, analogue noise, and quantisation noise), and the coupling and EMC-effects noise. The value of r is taken from the GPDB.

The uncertainty in the background measurement is

$$\sigma_{\mathrm{bg}}^2(k) = \frac{1}{n_{\mathrm{samp,bg}}}(n_{\mathrm{bg}}\tau + r^2), \tag{43}$$

here $n_{\mathrm{samp,bg}}$ is the number of sample used to measure the background. The number is set to 12, which is the number of pixels in the AC direction.

The calibration error σ_{cal} is calculated using concepts introduced in Jordi et al. (CJ-043). The idea is that the fluxes at some wavelengths will be better-calibrated than those at other wavelengths, because the instrument response is not constant. Thus at wavelengths with higher instrument responses, we can expect better calibrations than those with lower responses, and the uncertainty in calibration should thus be lower.

To internally calibrate the fluxes, we need to take into account the various effects that contribute to the formation of the spectrum. We can simply assume that the observed flux f_{λ_i} in a given

pixel index i with central wavelength λ_i is a linear combination of a few internally calibrated g_{λ_i} (Jordi et al., CJ-043):

$$f_{\lambda_i} = \sum_{j=-n}^{+n} a_{i,j} g_{\lambda_{(i+j)}}, \tag{44}$$

where $g_{\lambda_{(i+j)}}$ is the internally calibrated flux at the $(i+j)$-th pixel and $a_{i,j}$ are the transformation coefficients between the mean calibrated spectrum and the observed spectrum. The indices j correspond to the neighboring samples to the left and right of the sample being considered, which is the i-th pixel. Thus: $j = 0$ is the sample being considered, i.e. the i-th pixel. $j = -1$ is the sample on the left of the i-th pixel, $j = +1$ is the sample on the right, and so forth. The optimum number of the neighboring pixel n to be taken into account is at the moment being investigated (Jordi et al., CJ-043), and the convention at the moment is to take $n = 2$.

The uncertainties of the calibration coefficients $a_{i,j}$ will then translate into the uncertainty in calibration as

$$\sigma^2_{\mathrm{cal}_{\lambda_i}} = \sum_{j=-n}^{j=+n} \sigma^2_{a_{i,j}} g^2_{\lambda_{(i+j)}}. \tag{45}$$

The uncertainties in the calibration coefficients $a_{i,j}$ will depend ultimately on the measured fluxes f_{λ_i}. The lower the overall instrument response, the less we will measure any flux f_{λ_i}, and thus the uncertainy will be lower. The uncertainty in calibration coefficient is thus proposed to be of the form (Jordi et al., CJ-043)

$$\sigma_{a_{i,j}} = A + \frac{B}{C(\lambda_i)}, \tag{46}$$

here A and B are constants taken to be

$$A = \frac{0.001}{\sqrt{5}}, \tag{47}$$

$$B = \frac{0.0001}{\sqrt{5}}, \tag{48}$$

here the values are taken such that the maximum calibration error is about 1 mmag. The value $\sqrt{5}$ is taken because 5 samples are used for the calibration ($n = 2$). The function $C(\lambda_i)$ is defined to be

$$C(\lambda_i) = (D \times H) T_m(\lambda_i) T_p(\lambda_i) Q(\lambda_i), \tag{49}$$

which is the overall response of the instrument (mirrors, prism, and CCD) multiplied by the pupil area. To avoid an undefined values when $C(\lambda_i) = 0$, a threshold could be forced upon:

$$C(\lambda_i) = 0.01, \text{ for } C(\lambda_i) \leq 0.01. \tag{50}$$

8 Comparisons with GOG

In this section the results of Ulysses are compared with GOG. Some care has been taken to ensure that both software use the same version of the MDB and hence the same parameters for the instrument modelling. Special access to see the source code of GOG was granted by CU2, allowing us to study the inner workings of GOG, and to modify the code to run certain tests.

The comparisons reported in this section covers the comparison of BP/RP spectra from a flat input spectrum, the comparison of BP/RP spectra from various input spectra, and the signal to noise (SNR) comparison.

8.1 Flat spectra comparison

In these comparisons, the input spectrum is a flat spectrum (i.e. constant number of photons $\mathrm{nm}^{-1}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$). For all simulation runs, unless otherwise noted, only one set of LSFs is used: Those for the chip row number 4 and telescope number 1 (henceforth the S1R4T1 LSF). Using a flat input spectrum and simulating only a specific chip is useful to simplify the situation and reduce the contributions that will add up to the final, sampled spectrum.

We first compare the result using an averaged LSF for Ulysses and the range of integration as described in Section 3.3. The comparison is shown in Figure 16. By inspecting the top part of the Figure, we can see through a cursory glance that Ulysses and GOG are largely consistent with each other. However, by inspecting the residuals in the middle and bottom plots, we can see that there are small differences as a function of samples and wavelengths. At the bottom plots of Figure 16, the difference between GOG and Ulysses flux in a pixel is described in terms of the absolute relative difference (ARD), i.e.

$$\mathrm{ARD}_i = \frac{|\mathrm{Flux}_{\mathtt{GOG},i} - \mathrm{Flux}_{\mathtt{Ulysses},i}|}{\mathrm{Flux}_{\mathtt{GOG},i}}, \tag{51}$$

here i is the index of the pixel being considered.

The distribution of the ARDs at the bottom plots of Figure 16 is shown in Figure 17. These distributions of the ARDs for all pixels of a spectrum can be characterized in terms of the root mean square (RMS), the mean absolute relative difference (MARD), and the median absolute relative difference (MedARD). The RMS is the square-root of the arithmetic mean of all the absolute relative differences in the pixels:

$$\mathrm{RMS} = \sqrt{\frac{1}{n_{\mathrm{pix}}} \sum_{i=0}^{n_{\mathrm{pix}}-1} (\mathrm{ARD}_i)^2} \tag{52}$$

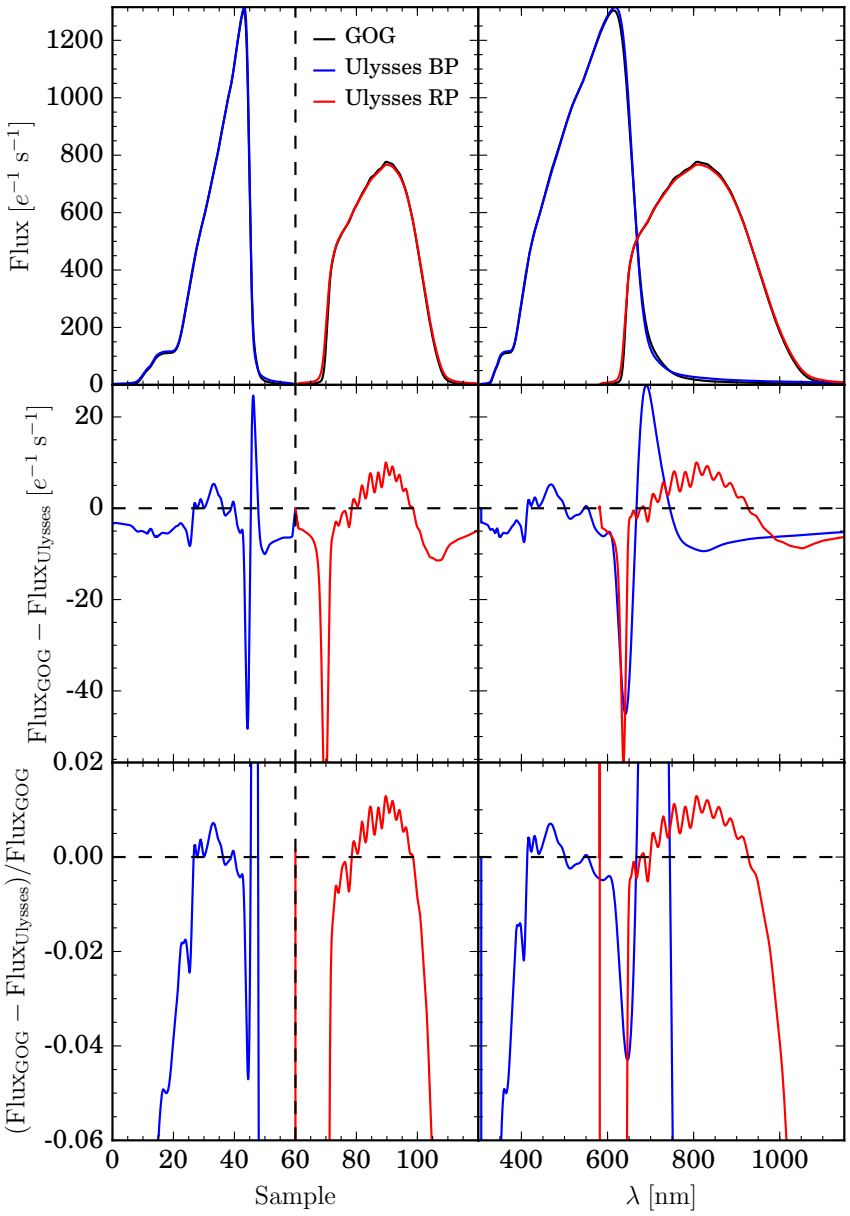


Figure 16: The comparison of a flat spectrum as simulated by `GOG` and `Ulysses`. The left columns plot the values labelled in the y-axes as a function of pixel indices in the spectrum, while the right columns plot the values as a function of wavelength. The top row plots show the BP/RP spectra as observed by `GOG` (solid black lines) and `Ulysses` (blue line for BP, red line for RP). The middle row plots show the residuals between the two calculations, while the bottom row plots show the relative difference between the two. The vertical dashed lines on the left columns indicate the termination of the BP spectra and the beginning of the RP spectra, while the horizontal dashed lines on the middle and the bottom rows indicate a perfect match between the two spectra as $y = 0$.

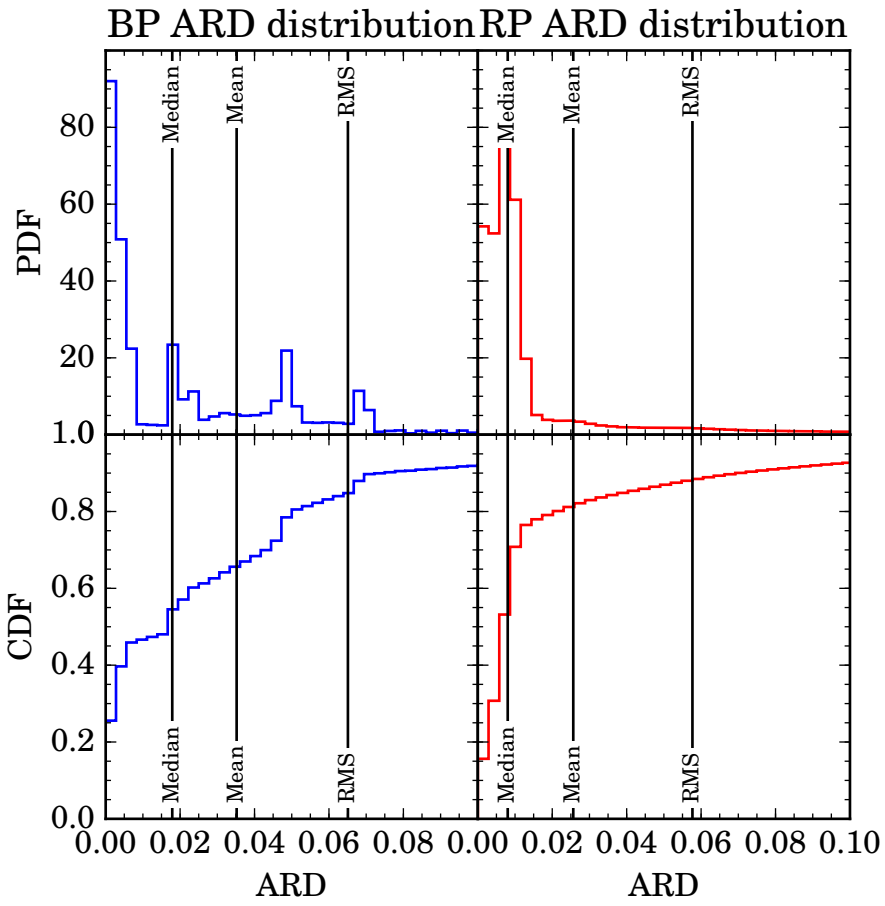


Figure 17: The distribution of BP ARD (left column) and RP ARD (right column). The top rows show the probability density function (PDF) of the BP/RP ARDs, while the bottom rows show the cumulative distribution function (CDF) of the BP/RP ARDs. The distribution of the ARDs can be characterized in terms of its root mean square (RMS), its mean, or the median. These values are indicated by the black vertical lines for both BP and RP.

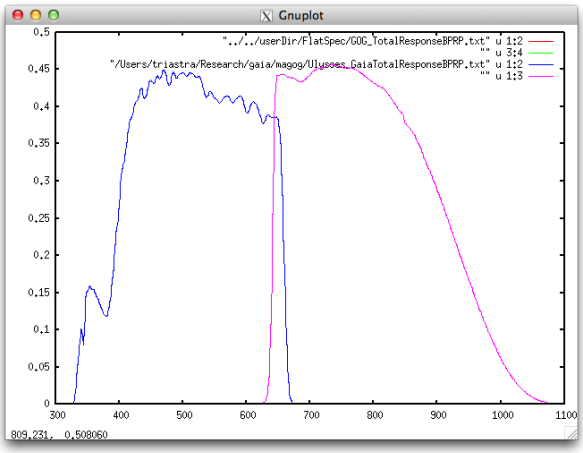
MARD is an average of all the absolute relative differences in the pixels:

$$\mathrm{MARD} = \frac{1}{n_{\mathrm{pix}}} \sum_{i=0}^{n_{\mathrm{pix}}-1} \mathrm{ARD}_i, \tag{53}$$

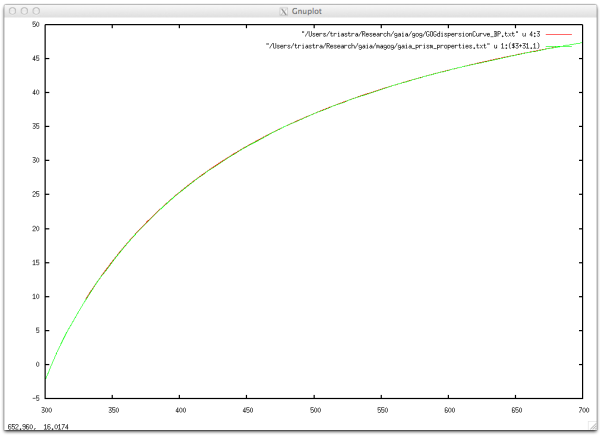
while MedARD is the median of the distribution of the absolute relative differences:

$$\mathrm{MedARD} = Q\left[0.5\,|\mathrm{ARD}\right], \tag{54}$$

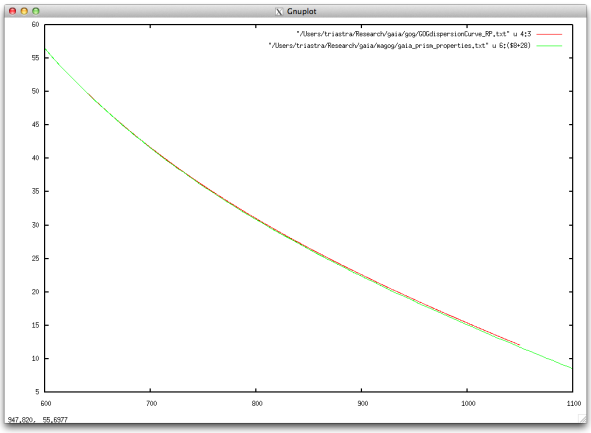
where Q is the quantile function (i.e. the inverse cumulative distribution function) of the absolute relative difference distribution.



(a) BP/RP total transmissivity curve



(b) BP dispersion curve



(c) RP dispersion curve

Figure 18: A GOG v. Ulysses comparison of three parameters: The total BP and RP transmissivity curve (top), the BP dispersion curve (middle), and the RP dispersion curve.

We are not interested in the edges of the spectra, since at the edges the transmissivity is close to zero, and this can seriously skew the distribution of the absolute relative differences (aside from that, the signals are also very low and in reality contain mostly instrumental noise). Thus in calculating RMS, MARD, and MedARD we do not include the pixel at the edges. We only include pixels corresponding to the wavelength range $330\ \mathrm{nm} \leq \lambda \leq 680\ \mathrm{nm}$ in BP and $640\ \mathrm{nm} \leq \lambda \leq 1050\ \mathrm{nm}$ in RP.

Thus, the MARD (MedARD) of Ulysses and GOG spectra in Figure 16 are 3.52% (1.79%) in BP and 2.56% (0.80%) in RP.

In finding out the sources of these differences, we first convince ourselves that both Ulysses

and GOG use the same parameters to model the dispersion curve, the transmissivity curve of the BP/RP spectrograph system, and the LSF. In Figure 18a, we see that the response curves employed by both GOG and Ulysses are the same.

To model the dispersion curves of the BP/RP prisms, GOG uses the fitted dispersion curve provided by Domingues & Rebordao (CDO-001). This curve is computed using the ray tracing software Zemax which model the optical system of *Gaia*. The resulting curve is then fitted using a bivariate polynomial of the 3rd order. The coefficients of this polynomial are stored in the GPDB. On the other hand, as described in Section 3.2, Ulysses calculates the dispersion curve and its inverse from first principle. The difference between GOG and Ulysses dispersion curves can be seen in Figure 18b and 18c. We see that the differences between the two are insignificant and they are identical in many respects.

Having convinced ourselves that both GOG and Ulysses use (more-or-less) the same values for the basic modelling, we proceed to investigate the difference between the two convolution and sampling algorithms. For this analysis, GOG and Ulysses are modified so that all other effects other than the spectral convolution and sampling are removed. This includes removing the transmissivity effects and using a flat input spectrum. The LSF used for both simulations are the S1R4T1 LSF.

The first result can be seen in Figure 19. We see that the BP/RP spectra as calculated by GOG and Ulysses are mostly identical, except at the edges of the spectra: For BP the divergence starts at the red edge from $\lambda \simeq 600$ nm and peaks at $\lambda \simeq 675$ nm, while for RP it starts at the blue edge from $\lambda \simeq 675$ nm and peaks at $\lambda \simeq 625$ nm. Here the MARD (MedARD) increases to 4.08% (1.93%) in BP and 3.44% (0.63%) in RP.

From the inspection of GOG's source code, we deduce that the main cause of the divergence in GOG and Ulysses spectra is the integration range of the LSF convolution. GOG integrates only from $\lambda = 330$ nm to $\lambda = 680$ nm for BP and from $\lambda = 640$ nm to $\lambda = 1050.2$ nm in RP, whereas Ulysses integrates from $\lambda = 325$ nm to $\lambda = 700$ nm for BP and from $\lambda = 625$ nm to $\lambda = 1050$ nm for RP. In addition to that, GOG's nominal LSF grids do not include $\lambda = 700$ nm in BP and $\lambda = 625$ nm in RP, which are available in the LSF analytical library. These differences in the integration ranges explain the extra fluxes at the edges of the spectra: There will be more fluxes at those wavelengths because of the additional wavelengths taken into account. That GOG writers decided to reduce the wavelength integration ranges, this is possibly because it is thought at these wavelength ranges the transmissivity of the instruments is nearly zero and does not contribute much to the shape of the final spectrum.

We run Ulysses again using these reduced integration limit. The results are shown Figure 20. We can see that the differences have been reduced considerably, with a MARD (MedARD) of 1.39% (1.65%) in BP and 0.96% (0.45%) in RP.

GOG also does not interpolate LSF for wavelengths in-between the nominal wavelength grid. For

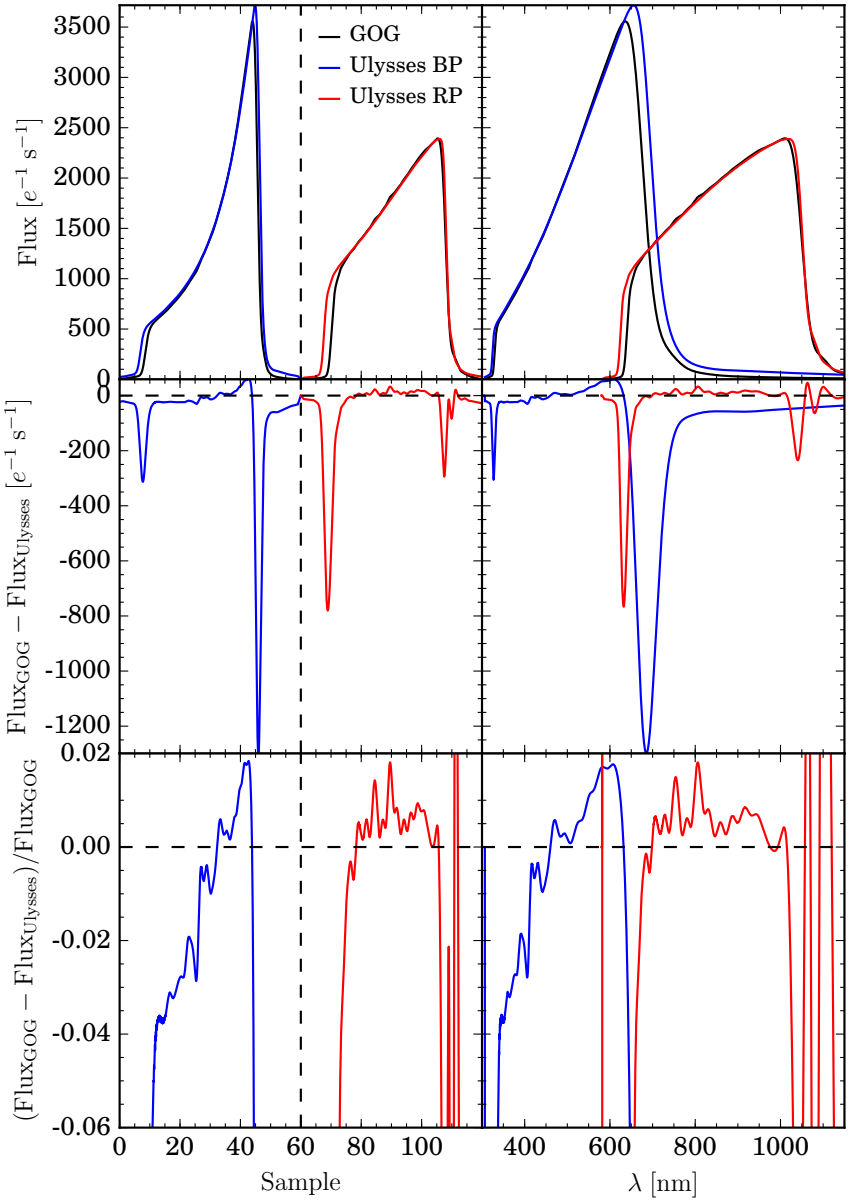


Figure 19: As in Figure 16, but with the transmissivity curves removed. Discrepancies start to appear in the spectra: For BP the difference between GOG and Ulysses starts at the red edge from $\lambda \sim 600$ nm and peaks at $\lambda \sim 675$ nm. For RP it starts at the blue edge from $\lambda \sim 675$ nm and peaks at $\lambda \sim 625$ nm.

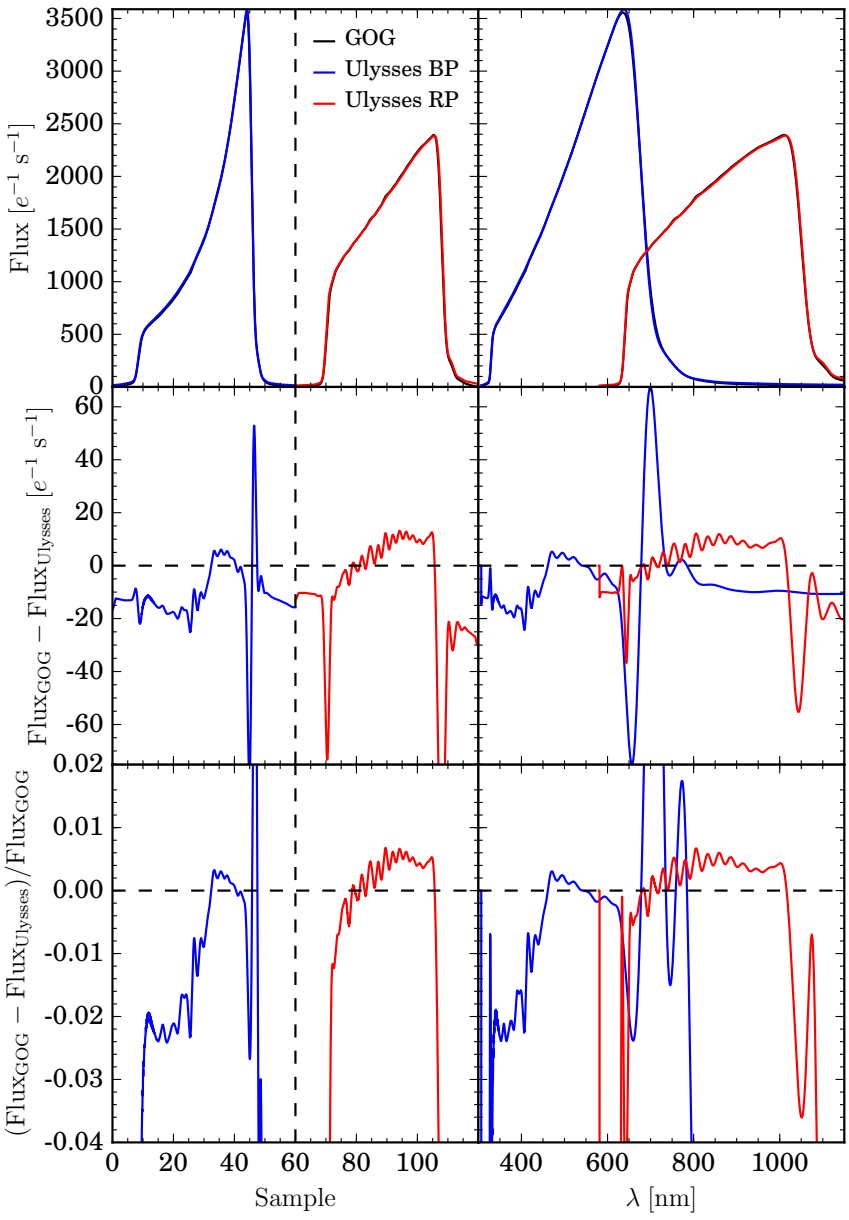


Figure 20: The same as in Figure 19, but with a reduced integration range for the LSF convolution in Ulysses. The LSF integration range is reduced such that it is similar to GOG: $\lambda = 330$ nm to $\lambda = 680$ nm for BP and from $\lambda = 640$ nm to $\lambda = 1050.2$ nm in RP.

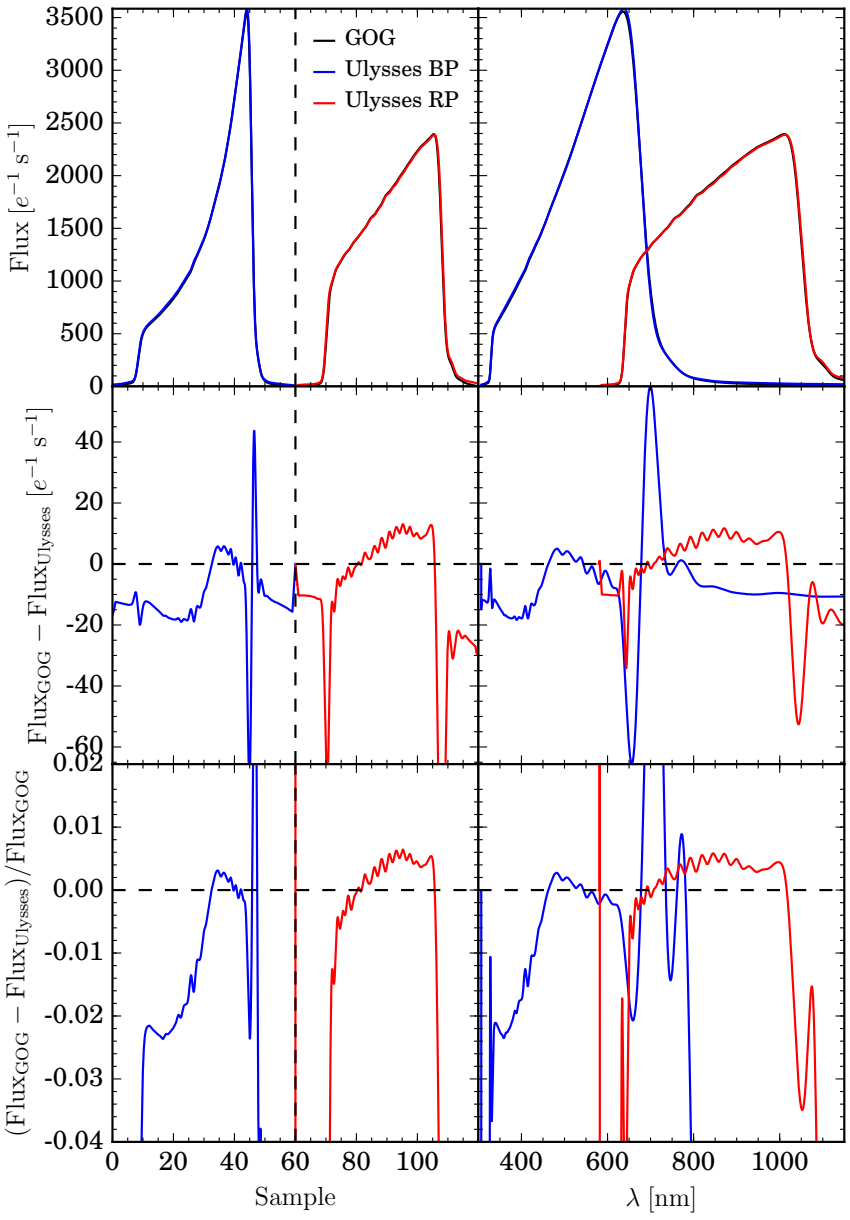


Figure 21: The same as in Figure 19, but with the nearest-neighbor integration scheme in addition to a reduced integration range for the LSF convolution in `Ulysses`.

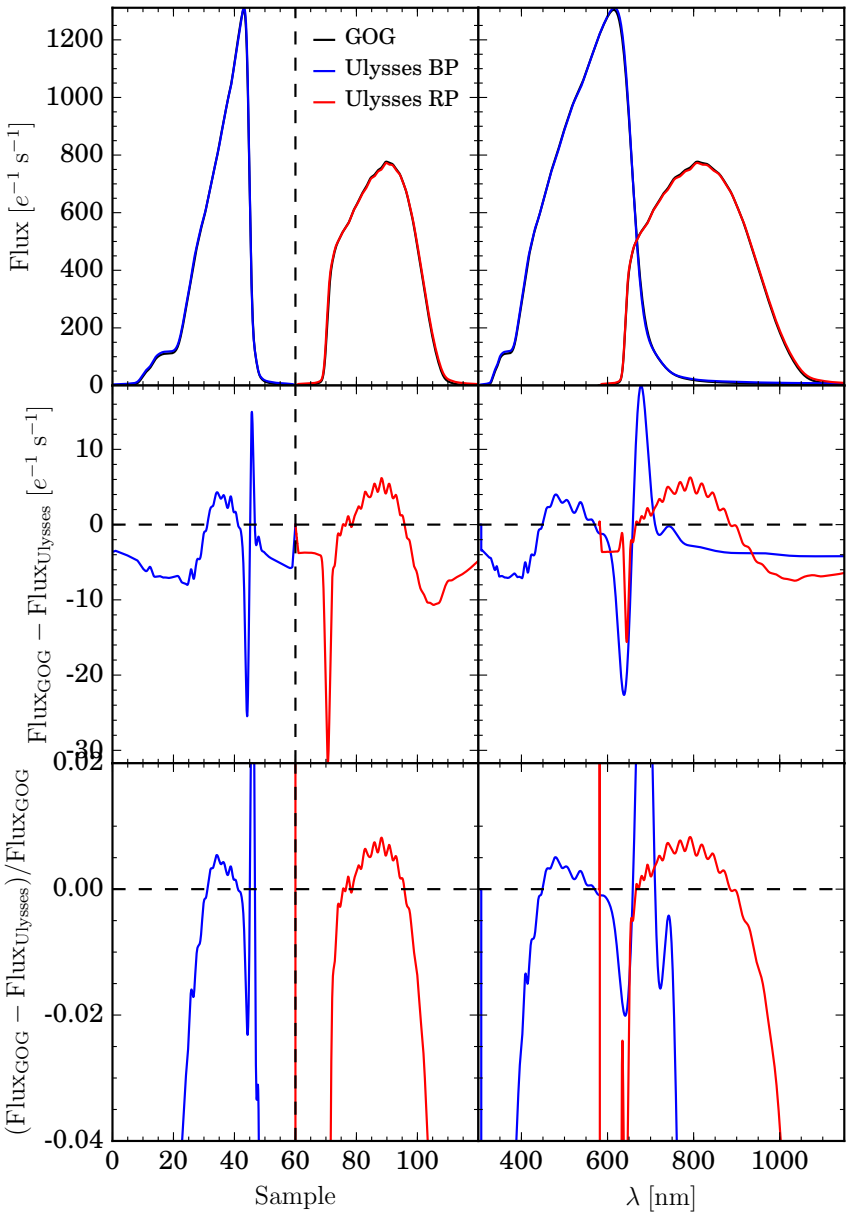


Figure 22: The same as in Figure 16, i.e. a flat spectrum convolved with *Gaia* transmissivity filter and the LSF, but with the nearest-neighbor integration scheme in addition to a reduced integration range for the LSF convolution in `Ulysses`.

the integration, GOG selects the nearest nominal wavelength in the grid. This "nearest-neighbor" integration scheme causes a discontinuity in the spectrum shape by at most 0.5% (Sartoretti & Isasi, PS-011). We imitate this integration by running Ulysses with the nearest-neighbor integration scheme as in GOG, and with the same integration range as well. The results are shown in Figure 21. Using this convolution scheme the MARD (MedARD) is further reduced a bit to 1.36% (1.65%) in BP and 0.94% (0.44%) in RP.

This modifications of Ulysses to imitate GOG in the integration range, integration scheme, and the usage of the LSF, however, do not necessarily reduce the differences between GOG and Ulysses spectra when the instrument profiles are re-applied. We simulate again the flat spectrum using Ulysses, using the methods described in the previous paragraph and re-applying the instrument profiles of the BP/RP system. The results are shown in Figure 22. Using this method, the MARD (MedARD) of the RP spectrum dropped only slightly to 2.21% (0.60%), however the MARD of the BP rises to 4.2% (although BP MedARD is reduced slightly to 1.70%).

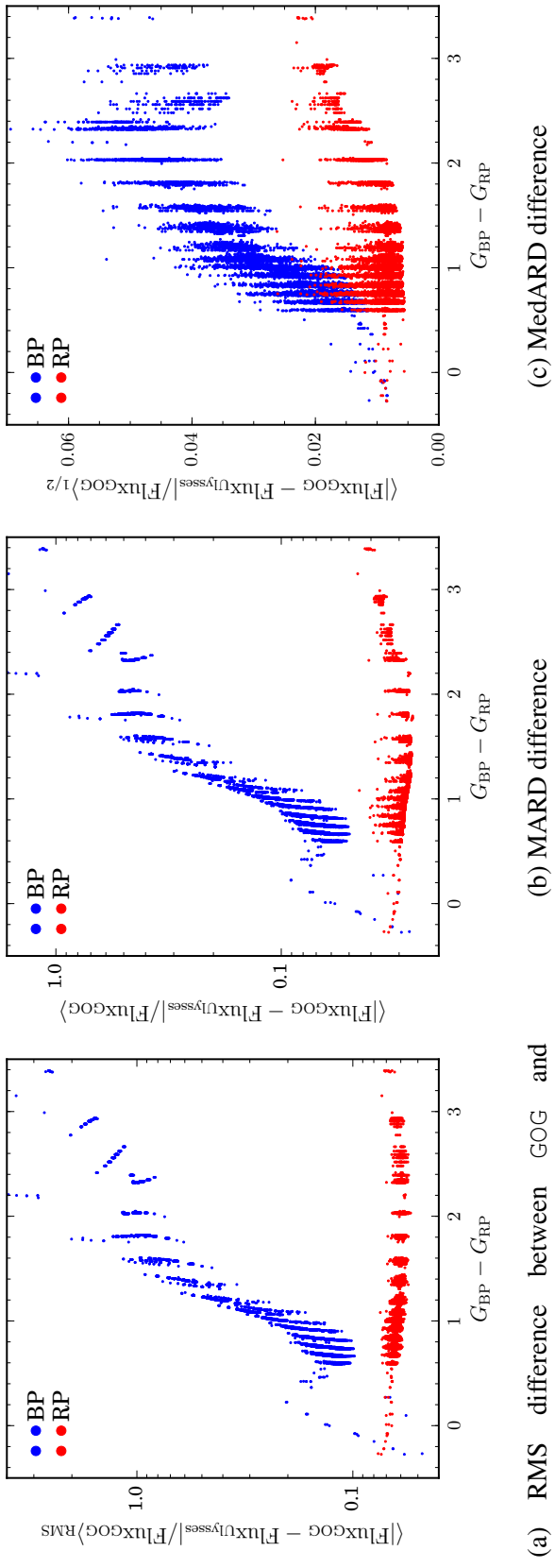
From this exercise we can conclude that for (nearly) flat spectra, the maximum differences between GOG and Ulysses spectra—in terms of MARD–will be up to $\sim 4\%$, if Ulysses is set to convolve the input spectra with an averaged LSF and the full integration range. A summary of all the differences for different Ulysses convolution methods described in this Section is given in Table 4.

In summary, the differences in the convolution method employed by GOG and Ulysses can result in markedly different spectra. The differences in the convolution method are as follows:

- The integration interval in Ulysses is longer than GOG. Ulysses integrates the wavelengths in the BP regime between $325\ \mathrm{nm} \leq \lambda \leq 700\ \mathrm{nm}$ and $625\ \mathrm{nm} \leq \lambda \leq 1050\ \mathrm{nm}$ in the RP regime, while GOG integrates in the interval of $330\ \mathrm{nm} \leq \lambda \leq 625\ \mathrm{nm}$ for BP and $640\ \mathrm{nm} \leq \lambda \leq 1050.2\ \mathrm{nm}$ for RP.
- Ulysses uses an additional wavelength point for the LSF nominal grid, at $\lambda = 700\ \mathrm{nm}$ in BP and $\lambda = 625\ \mathrm{nm}$ in RP. GOG uses the LSF nominal grid only up to $\lambda = 675\ \mathrm{nm}$ in BP and $\lambda = 650\ \mathrm{nm}$ in RP. The additional information allows us to calculate in more precise the additional light contribution to each pixel from these wavelength.
- Ulysses uses bicubic spline interpolation to interpolate the LSF outside the nominal wavelengths, while for the same points GOG uses the nearest LSF (i.e. a constant interpolation).

Table 4: A summary of the differences between GOG and Ulysses in terms of the root mean square differences (RMS), mean absolute relative differences (MARD), and the median of the absolute relative differences (MedARD). The input spectrum for all cases of convolution is a flat spectrum.

Transmissivity			Figure	BP			RP		
				RMS	MARD	MedARD	RMS	MARD	MedARD
On	Averaged, full integration range		16	0.065	0.035	0.018	0.058	0.026	0.008
	S1R4T1 nearest-neighbor, reduced integration range		22	0.073	0.042	0.017	0.045	0.022	0.006
Off	S1R4T1 interpolated, full integration range		19	0.098	0.041	0.019	0.115	0.034	0.006
	S1R4T1 interpolated, reduced integration range		20	0.018	0.014	0.016	0.023	0.010	0.005
	S1R4T1 nearest-neighbor, reduced integration range		21	0.018	0.014	0.015	0.022	0.009	0.004



(a) RMS difference between GOG and Ulysses spectra

(b) MARD difference

(c) MedARD difference

Figure 23: The comparison between GOG and Ulysses spectra, as a function of their $(G_{BP} - G_{RP})$ color. Here the differences between the spectra are represented by the root mean square of the absolute relative differences on the left, the mean absolute relative differences (MARD) in the middle, and the median absolute relative differences (MedARD) on the right, of BP and RP spectra. We can observe here a consistent increase of MARD and MedARD for BP spectra with increasing red color, while the MARD and MedARD for RP spectra are (almost) always constant.

8.2 Comparisons of GOG and Ulysses noise-free spectra

To analyse the differences between GOG and Ulysses spectra for various kinds of spectra, BP/RP spectra have been simulated on a common set of stellar spectra—produced by GOG using the Besançon model and a set of spectral libraries—using both GOG and Ulysses. These common set are divided into two cases: Those with varying magnitudes ranging from $G \sim 8$ to $G \sim 20$, and those with constant magnitude at $G = 15$. The former case is used to analyse the absolute relative differences, while the latter is used to analyse the absolute residual.

In these cases, no noise were generated. We thus have two set of noise-free spectra for the same sources, processed by both GOG and Ulysses, and we can compare the differences between the two simulators.

8.2.1 Sources with varying magnitudes

For all sources, we calculate the RMS, MARD, and MedARD of both BP and RP spectra. The plot of these three values as a function of color is shown in Figure 23. We can see that the redder a star is, the more pronounced are the differences between GOG and Ulysses. Redder stars are cooler and thus have more absorption lines, making the input spectrum more uneven and very complex. The convolution of the input spectrum with the LSF is basically a calculation of the contribution of fluxes from all wavelength on a given point on the focal plane. Since the convolutions of GOG and Ulysses are, as previously discussed, markedly different (from the range of the wavelength in which Ulysses interval of integration is more extensive, the method of LSF interpolation in which Ulysses performs bicubic-spline interpolation while GOG uses the closest wavelength to the wavelength being considered), thus we can expect that the largest differences beetween GOG and Ulysses will be the spectra with the most complexity.

Figures 24–27 show the comparison of GOG and Ulysses BP/RP spectra for the best and the worst cases. Figure 24 shows the best match between GOG and Ulysses BP spectra, Figure 26 shows the best match in RP, while Figures 25 and 27 shows the worst match between GOG and Ulysses respectively for BP and RP. By comparing these BP/RP spectra with their input spectra we can see what sort of spectrum will give the best convergence and which will potentially diverges.

In Figure 28, the input spectra that provides the best and worst match between GOG and Ulysses are shown. Figure 28b indicate that cool stars with many absorption lines will potentially provide the largest BP difference between Ulysses and GOG, as indicated as well in Figure 23. For RP the complexity of the spectrum does not seem to be a problem.

It should be noted that the differences in term of the absolute relative differences of the spectra can also be caused by the fact that the flux in the spectra is already very low to begin with. Since in calculating the ARD we are dividing by the flux, a spectrum can have a high ARD because the

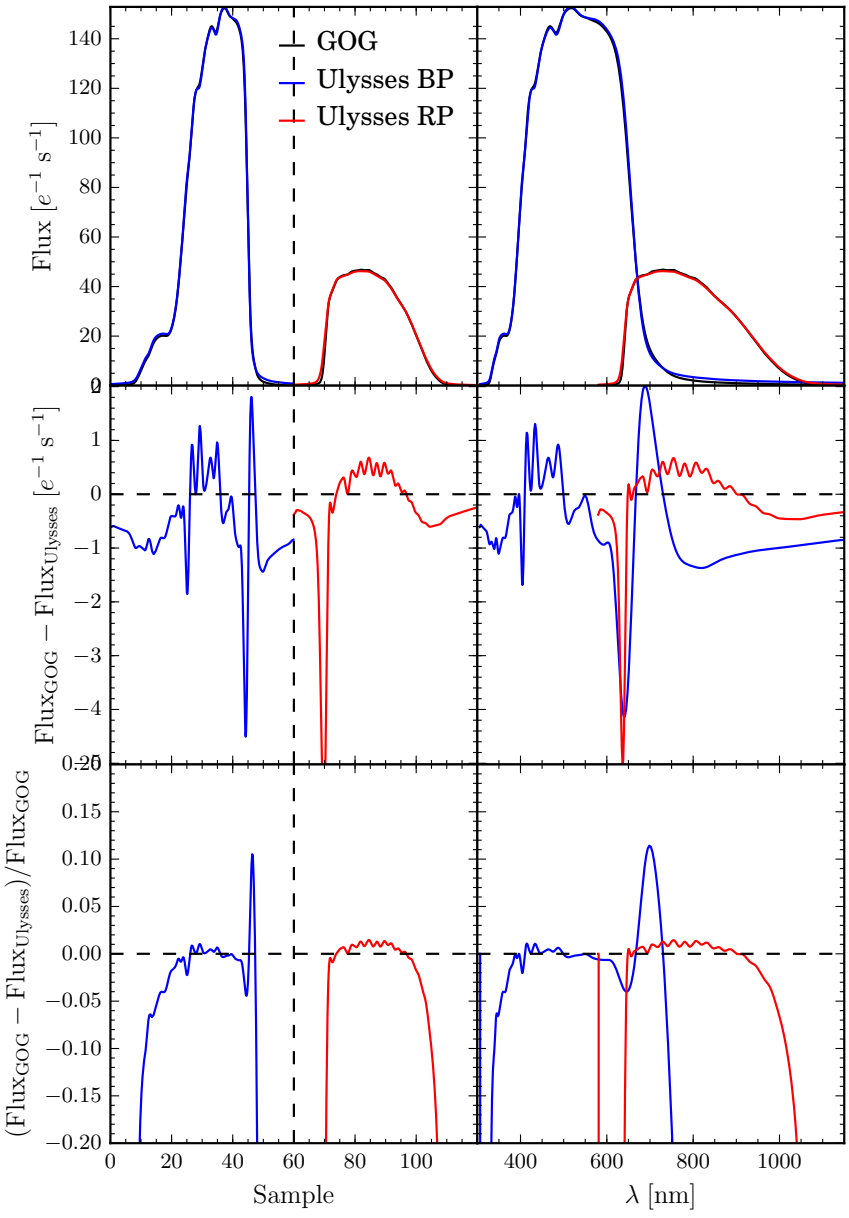


Figure 24: The same as in Figure 16, but for an input spectra with the lowest BP MedARD among the simulated spectra (i.e. the best BP match between GOG and Ulysses). This star has a MedARD of 0.83% in BP. The corresponding input spectrum that results in these spectra, along with their astrophysical parameters (AP) is shown in Figure 28a.

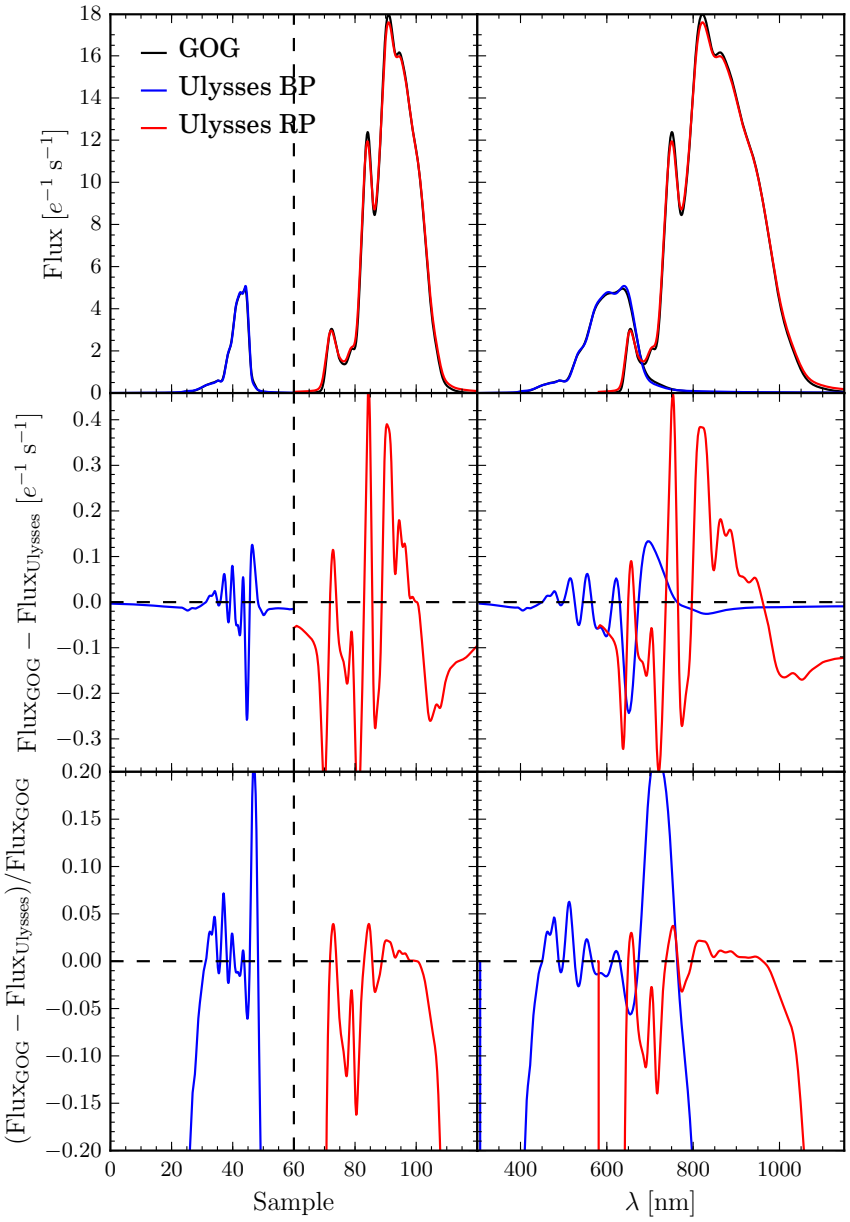


Figure 25: The same as in Figure 24, but for an input spectra with the highest BP MedARD among the simulated spectra (i.e. the worst BP match between GOG and Ulysses). This star has a BP MedARD of 13.27% in BP. The corresponding input spectrum that results in these spectra, along with their AP is shown in Figure 28b.

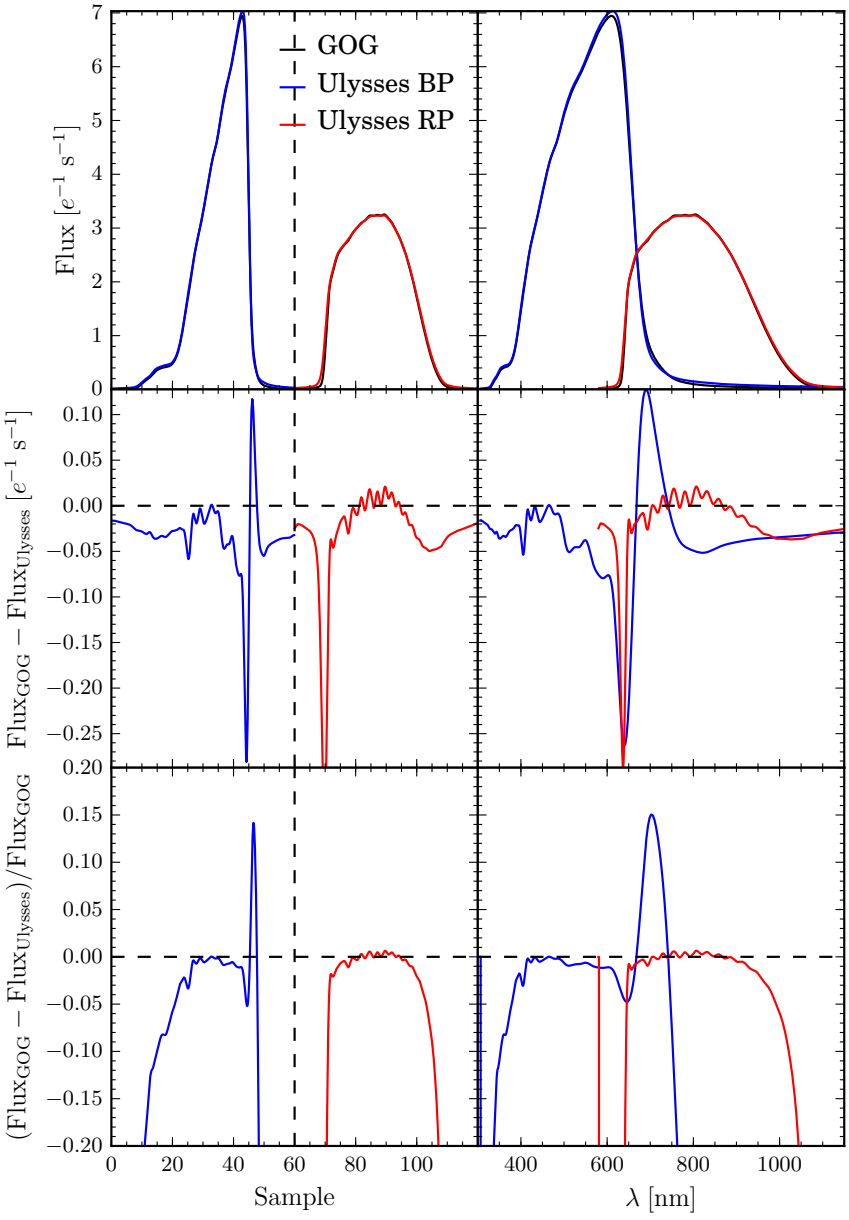


Figure 26: The same as in Figure 24, but for an input spectra with the lowest RP MedARD among the simulated spectra. This star has a MedARD of 0.56% in RP. The corresponding input spectrum that results in these spectra, along with their AP is shown in Figure 28c.

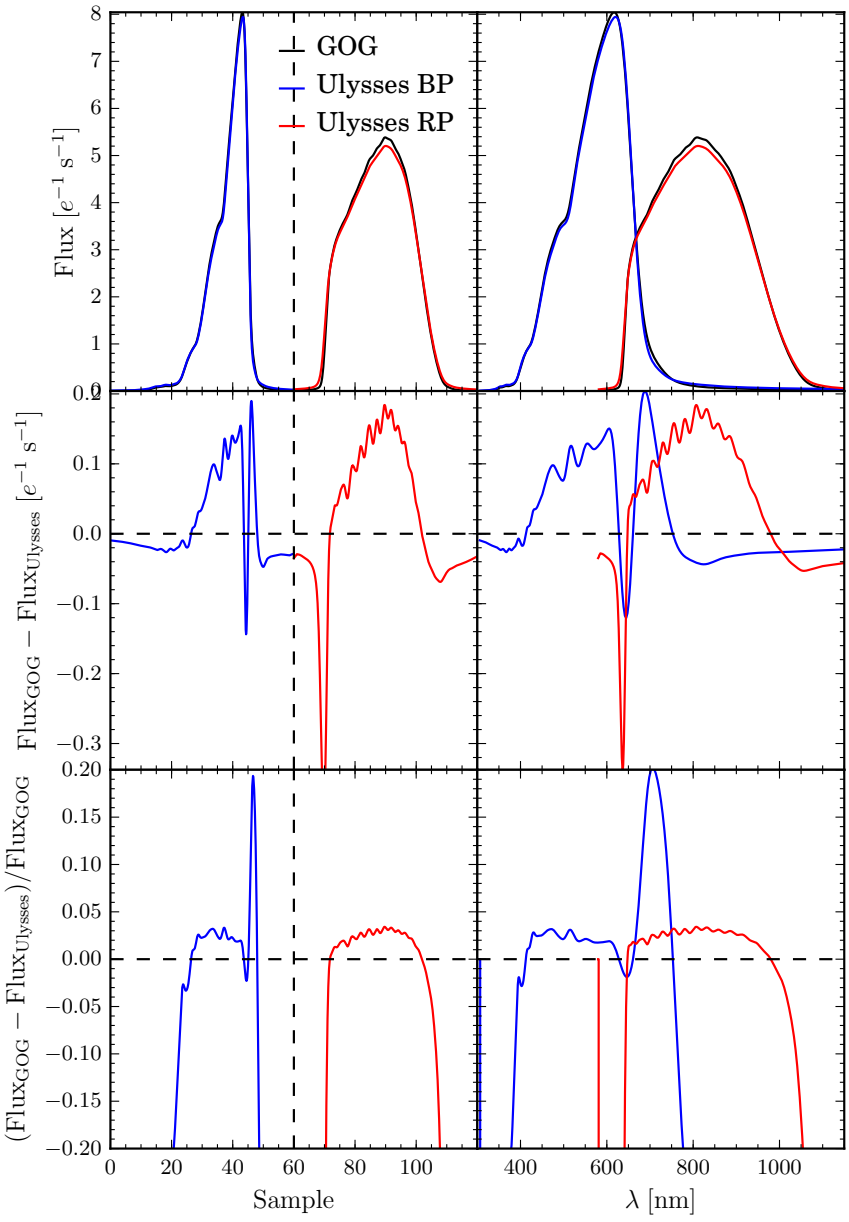
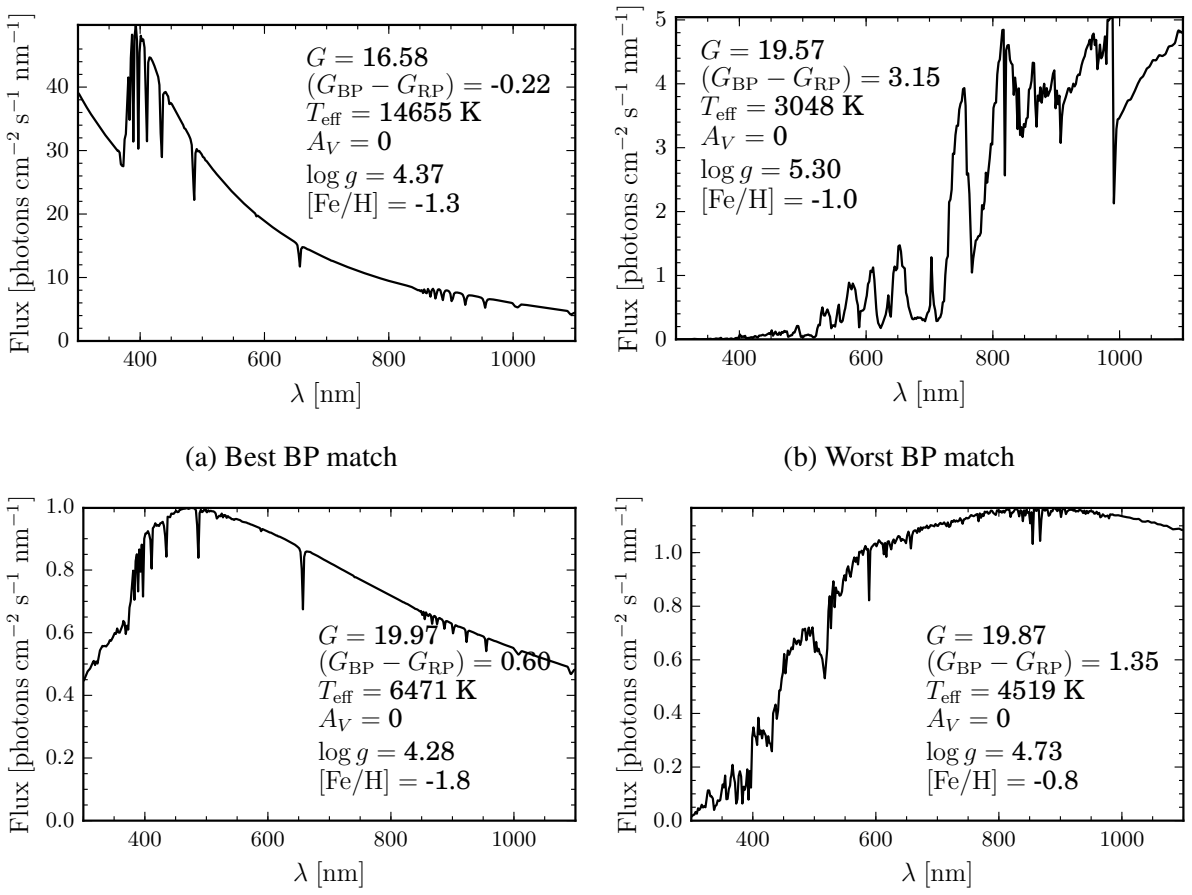


Figure 27: The same as in Figure 24, but for an input spectra with the highest RP MedARD among the simulated spectra. This star has a MedARD of 2.62% in RP. The corresponding input spectrum that results in these spectra, along with their AP is shown in Figure 28d.



(a) Best BP match

(b) Worst BP match

(c) Best RP match

(d) Worst RP match

Figure 28: The input spectra that produces the best and worst BP/RP spectra match between GOG and Ulysses, shown in Figures 24–27.

fluxes in the pixels are low. In order to disentangle this effect, in the following we thus analyse the differences between GOG and Ulysses in terms of their absolute residuals.

8.2.2 Sources with constant magnitudes at $G = 15$

For this analysis the common set with all sources having the same magnitude, at $G = 15$, is used. For all pixel in a spectrum, we can calculate the residual difference between GOG and Ulysses flux. We then take the absolute value, and for a spectrum with varying absolute residual as a function of wavelength we can calculate the mean absolute residuals (MAR), the median absolute residuals (MedAR), and its root mean square of absolute residuals. This is analogous to Equations 52–54.

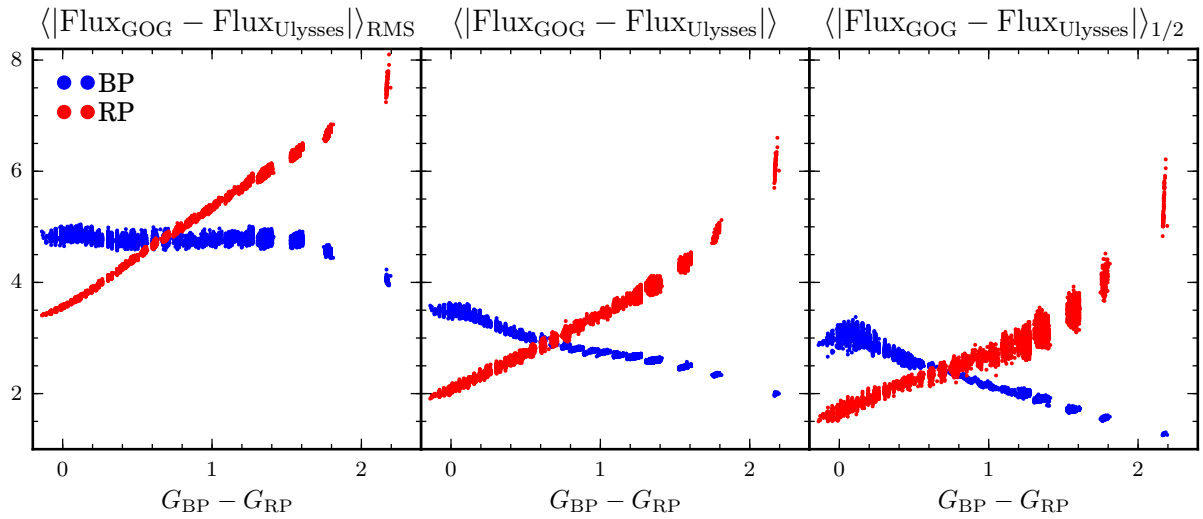


Figure 29: The comparison between GOG and Ulysses spectra, as a function of their $(G_{BP} - G_{RP})$ color. Here the differences between the spectra are represented by the root mean square of the absolute residual on the left, the mean absolute residual (MAR) in the middle, and the median absolute residual (MedAR) on the right, of BP and RP spectra. Here we can observe a trend inverse to what were observed in Figure 23: Here for example blue sources in fact have higher absolute residuals in BP than the red ones, simply because blue stars have higher fluxes than the red ones.

In Figure 29, the RMS of the absolute residuals, the mean absolute residuals, and the median absolute residuals (MedAR) are shown. We can observe a reversal of trend here: Blue-colored sources have in fact the highest AR in BP and red-colored sources have the lowest AR in BP. Whereas RP MedARD is relatively constant with regard to colors, here RP MedAR in fact increases with color. This can be explained by the fact that for red stars, there will be lower fluxes in the blue part of the spectrum and higher fluxes in the red part, thus lowering and increasing the residuals in their respective part. In absolute values, this increase and decrease in residual is not much: Between the highest and the lowest value, the median residual in BP only decreases by as much as $\sim 2\ e^-\ \mathrm{s}^{-1}$, the increase in RP by as much as $\sim 5\ e^-\ \mathrm{s}^{-1}$.

8.3 Error model comparison

Ulysses employs the same error model as GOG. They are described in Section 7. Thus we should not expect too many differences in the final calculation of the errors other than from the fact that the spectra themselves (which enter the noise model as described in Section 7) differs slightly, as described in the previous Subsection. Figure 30 shows the signal-to-noise ratio (SNR) of GOG and Ulysses as a function of the source G-band magnitude for a common set of sources. These sources are the same one generated in Section 8.2, but with the addition of noise in the spectra. In calculating the mean signal-to-noise ratio, only pixels in the wavelength

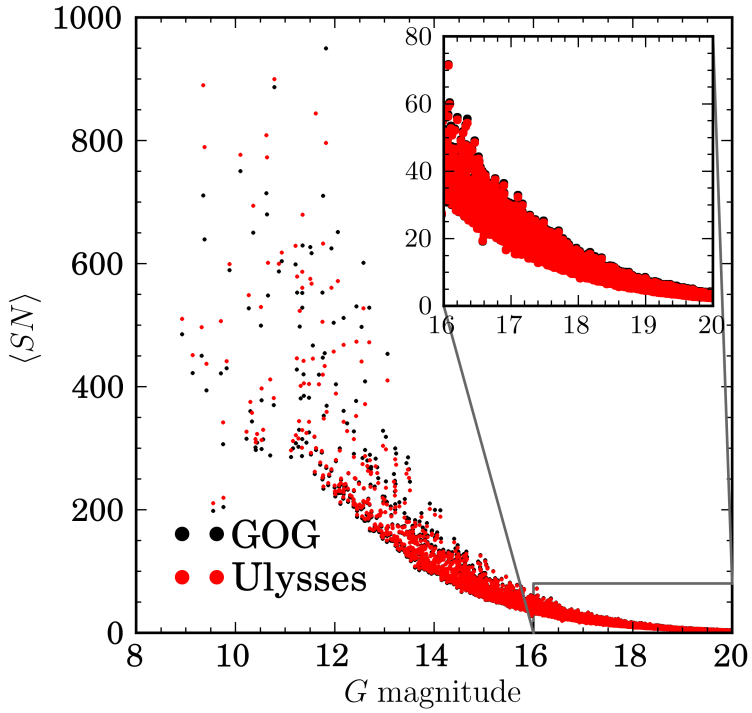


Figure 30: The comparison between GOG and Ulysses signal-to-noise ratio (SNR) as a function the G-band magnitude for each source. The sources are generated using GOG, and then simulated using both simulator. While the SNRs for sources fainter than $G \sim 12$ show good consistency, the SNRs for brighter sources diverge slightly. This is because the wavelength scale of GOG's sampled mean spectra has an additional uncertainty factor, which makes it difficult to make a selection based on wavelength.

range $330\ \mathrm{nm} \leq \lambda \leq 680\ \mathrm{nm}$ in BP and $640\ \mathrm{nm} \leq \lambda \leq 1050\ \mathrm{nm}$ in RP are included.

We can see that the SNRs are basically the same, except for sources brighter than $G \sim 12$. The latest version of Ulysses has the same gating scheme as GOG. The differences are thus interpreted as a consequence of excluding the pixels which wavelength lie outside the considered interval, while there is an artificially-added uncertainty in the calculation of GOG wavelength. Because GOG calculates the wavelength scales by adding uncertainties in the dispersion curve, GOG and Ulysses have different wavelength scale. Because of this added uncertainty, the resulting number of GOG pixels that lie in the given interval will be different with those of Ulysses. Consequently the total flux will be different. For fainter sources these additional or reduction of pixel fluxes are negligible, but for brighter sources the differences matter.

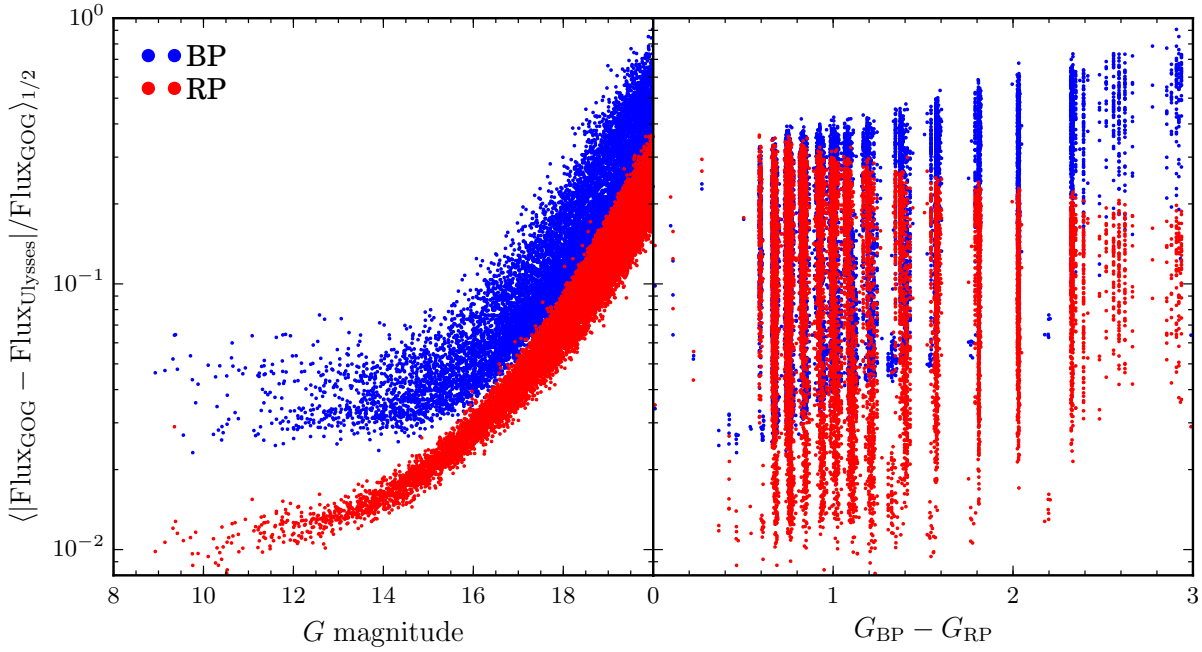


Figure 31: A comparison between GOG and Ulysses noisy spectra, here plotted as a function of the stars' G-band magnitude (left) and $G_{BP} - G_{RP}$ color (right). The differences in this plot are represented by the MedARD of each star. We can observe here a strong correlation between MedARD and G. The increase of MedARD with increasing red color for BP spectra is also observed here.

8.4 Comparisons of GOG and Ulysses noisy spectra

We can also compare the differences between GOG and Ulysses when noise is present. This exercise is especially useful when we are analyzing actual *Gaia* BP/RP spectra and comparing the models created by GOG and Ulysses. Knowing which area (e.g. in terms of a star's G-band magnitude and $G_{BP} - G_{RP}$ color) that will potentially presents the largest difference between GOG and Ulysses spectra will be useful.

For this exercise we use the same common set of sources already described in the previous Subsections, namely the set with noise added. In Figure 31 the MedARD for all stars as a function of the G-band magnitudes and $G_{BP} - G_{RP}$ is shown. As expected, the presence of increasing noise (which increase with as the star gets fainter) increase the MedARD. This effect is particularly apparent in BP. On the right plot of Figure 31, MedARD also increases as the star gets redder. This is because as $G_{BP} - G_{RP}$ increases, the BP part of the spectra gets fainter and thus the noise increase. The increasing presence of noise thus increase the differences between GOG and Ulysses.

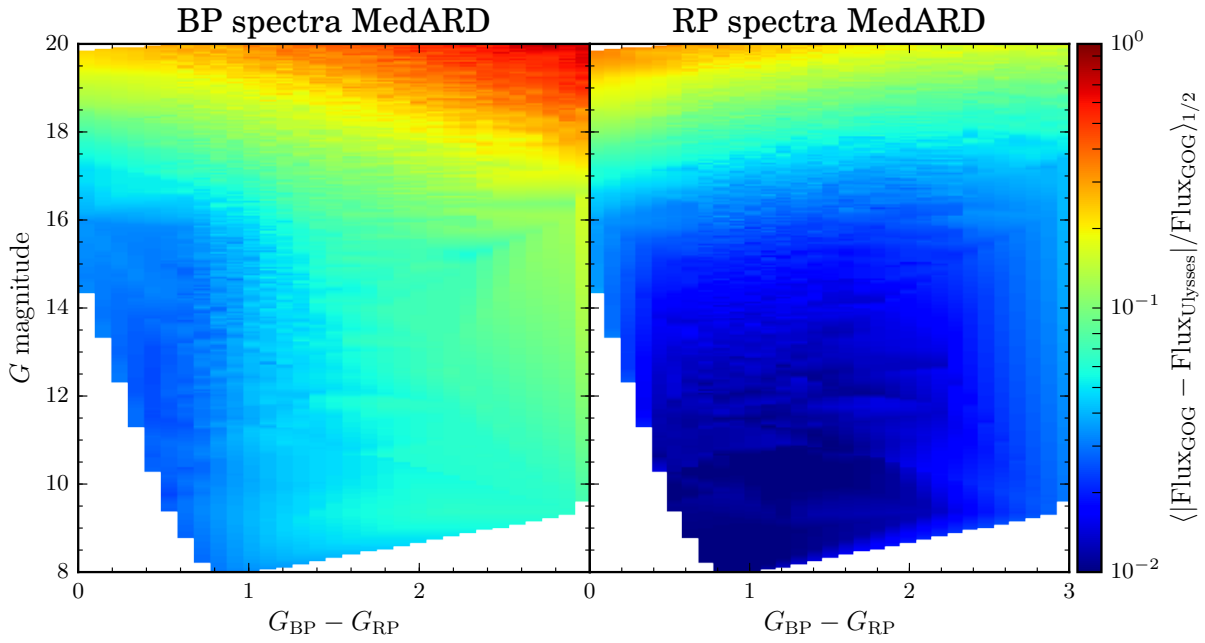


Figure 32: Plots of MedARD map in BP (left) and RP (right). These maps are made by linearly interpolating the scattered (in G and $G_{BP} - G_{RP}$ space) stellar data set into a regularly-spaced grid. The blank area is caused by the failure of the linear interpolation to calculate meaningful values due to sparse or nonexistent data in that particular area.

In Figure 32 the variations of BP and RP MedARD with G and $G_{BP} - G_{RP}$ are shown in two dimensional maps. Here we can see a different trend in BP and RP. For BP, the redder a star is, the greater the differences between GOG and Ulysses. Differences with $\mathrm{MedARD} \lesssim 5\%$ can only be found in stars with $G \lesssim 16$ and $G_{BP} - G_{RP} \lesssim 1$. In contrast, the differences in RP are generally smaller for a star with the same location in G-$G_{BP} - G_{RP}$ space. For stars with $G \lesssim 16.5$, RP MedARD is always $\lesssim 5\%$.

We also observe an opposite trends of MedARD in BP and RP. For BP, MedARD gets worse with increasing color. For RP, it is the opposite: for fainter stars ($G \gtrsim 17$) MedARD can actually get slightly better.

9 Closing

In this TN the main principles of Ulysses have been described, as well as its implementation in the code. The design and implementation of spectral convolution and sampling have been described, and its differences with respect to GOG have been documented. Readers should be aware that there will always be a discrepancy between Ulysses and GOG BP/RP spectra, and

the sources of these discrepancies have been identified.

The readers are now invited to test Ulysses itself, which can be checked out from the SVN repository: http://gaia.esac.esa.int/dpacsvn/DPAC/CU8/MPIA/software/ulysses/. Comments and corrections can be sent to the author of this TN through astraatmadja@mpia.de.

10 Acknowledgements

I am grateful to Anthony Brown for providing me the XpSim code, which was very helpful in the early development of Ulysses. I am also indebted to CU2 for providing me with the access to the GOG source code, without which a comparison between the two softwares would not have been possible.

References

[CBJ-066], Bailer-Jones, C., 2013, *Minutes of the eleventh CU8 meeting*, GAIA-C8-MN-MPIA-CBJ-066, URL http://www.rssd.esa.int/cs/livelink/open/3201815

Bailer-Jones, C.A.L., 2011, MNRAS, 411, 435, ADS Link

[AB-005], Brown, A., 2006, *Simulating prism spectra for the EADS-Astrium Gaia design*, GAIA-CA-TN-LEI-AB-005, URL http://www.rssd.esa.int/cs/livelink/open/550508

[AB-006], Brown, A., 2006, *Interface document for ad-hoc simulations of prism spectra for the EADS-Astrium Gaia design*, GAIA-C8-SP-LEI-AB-006, URL http://www.rssd.esa.int/cs/livelink/open/699404

[AB-009], Brown, A., 2007, *Photometry with dispersed images - overview of BP/RP data processing*, GAIA-C5-TN-LEI-AB-009, URL http://www.rssd.esa.int/cs/livelink/open/2329785

[AB-029], Brown, A.G.A., De Angeli, F., Evans, D.W., 2014, *Simulations of BP/RP data with XpSim*, GAIA-C5-TN-LEI-AB-029, URL http://www.rssd.esa.int/cs/livelink/open/3252246

[DB-007], Busonero, D., Gardiol, D., 2009, *LSF analytical library for simulations release 6.0*, GAIA-C3-TN-INAF-DB-007, URL http://www.rssd.esa.int/cs/livelink/open/2885858

Cardelli, J.A., Clayton, G.C., Mathis, J.S., 1989, ApJ, 345, 245, ADS Link

[CDO-001], Domingues, C., Rebordao, J., 2009, *Dispersion Analysis*, GAIA-C2-TN-INET-CDO-001, URL http://www.rssd.esa.int/cs/livelink/open/2860520

[GAIA.ASF.RP.SAT.00005], EADS Astrium, 2011, *Science Performance Budget Report*, GAIA.ASF.RP.SAT.00005, URL http://www.rssd.esa.int/cs/livelink/open/2722557

Fitzpatrick, E.L., 1999, PASP, 111, 63, ADS Link

[DG-011], Gardiol, D., Busonero, D., Corcione, L., et al., 2009, *PSF/LSF model for simulation purposes*, GAIA-C2-TN-INAF-DG-011, URL http://www.rssd.esa.int/cs/livelink/open/2879435

Hecht, E., 2001, *Optics 4th edition*

Jordi, C., Høg, E., Brown, A.G.A., et al., 2006, MNRAS, 367, 290, ADS Link

[CJ-043], Jordi, C., Fabricius, C., Carrasco, J.M., et al., 2008, *Error model for the mean BP/RP spectra for GOG cycle 4 simulations*, GAIA-C5-TN-UB-CJ-043, URL http://www.rssd.esa.int/cs/livelink/open/2816701

Jordi, C., Gebran, M., Carrasco, J.M., et al., 2010, A&A, 523, A48, ADS Link

[GAIA-LL-046], Lindegren, L., 2003, *Representationof LSF and PSF for GDAAS-2*, GAIA-LL-046, URL http://www.rssd.esa.int/cs/livelink/open/357835

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992, *Numerical recipes in C. The art of scientific computing*

[PS-011], Sartoretti, P., Isasi, Y., 2011, *Spectra Simulation in GOG*, GAIA-C2-TN-OPM-PS-011, URL http://www.rssd.esa.int/cs/livelink/open/3056260

A Acronyms used in this TN

The following table has been generated from the on-line Gaia acronym list:

Acronym	Description
AC	ACross scan (direction)
ADC	Analogue-to-Digital Converter
AF	Astrometric Field (in Astro)
AL	ALong scan (direction)
AP	Astrophysical Parameter
AR	Anti-Reflection coating (CCD)
ARD	Absolute Relative Difference
BP	Blue Photometer
CCD	Charge-Coupled Device
CDF	Calibration Data File
CTI	Charge Transfer Inefficiency
DPAC	Data Processing and Analysis Consortium
FITS	Flexible Image Transport System
FOV	Field of View (also denoted FOV)
GPDB	Gaia Parameter DataBase
LSF	Line Spread Function
MAR	Mean Absolute Residual
MARD	Mean Absolute Relative Difference
MDB	Main DataBase
MedARD	Median Absolute Relative Difference
PDF	Probability Density Function
PSF	Point Spread Function
PhotPipe	Photometric Pipeline (CU5 / DPCI)
RMS	Root Mean Square
RON	Read-Out Noise (CCD)
RP	Red Photometer
RVS	Radial Velocity Spectrometer
SED	Spectral Energy Distribution
SM	Sky Mapper
SNR	Signal-to-Noise Ratio (also denoted SN and S/N)
SVN	SubVersioN
TDI	Time-Delayed Integration (CCD)
TN	Technical Note
UV	UltraViolet

B Converting between flux densities

The input spectra of Ulysses must be in units of nanometer for wavelength and photons cm^{-2} s^{-1} nm^{-1} for flux density. In many cases, spectral libraries are not written in these units. Wavelengths are sometimes written in Ångstroms (Å), and flux densities are written in erg cm^{-2} s^{-1} Å^{-1}, or erg cm^{-2} s^{-1} Hz^{-1}. At the moment Ulysses provides several unit conversions. The following are the derivations of the flux unit conversions handled by Ulysses, which can also serves as handy guides in converting wavelengths and flux densities.

B.1 From erg cm^{-2} s^{-1} Å^{-1} to photons cm^{-2} s^{-1} nm^{-1}

To convert from Ångstrom to nanometer, we only need to remember that $1\ \text{Å} = 10^{-10}$ m. Thus $1\ \text{Å} = 10^{-1}$ nm.

The conversion of flux densities from erg cm^{-2} s^{-1} to photons nm^{-1} cm^{-2} s^{-1} can be done by relating energy to number of photons through

$$E = n\frac{hc}{\lambda}, \tag{55}$$

where h is the Planck constant and c is the speed of light. We then have

$$F_\lambda \left[\frac{\text{erg}}{\text{cm}^2\ \text{s Å}}\right] = \frac{dE}{d\lambda} \tag{56}$$

$$= \frac{hc}{\lambda}\frac{dn}{d\lambda}\left[\frac{\text{photon}}{\text{cm}^2\ \text{s Å}}\right], \tag{57}$$

$$f_\lambda \left[\frac{\text{photons}}{\text{cm}^2\ \text{s Å}}\right] = \frac{1}{hc}\lambda F_\lambda \left[\frac{\text{erg}}{\text{cm}^2\ \text{s Å}}\right]. \tag{58}$$

The value of hc is 1.986×10^{-25} J m, or $hc = 1.986 \times 10^{-8}$ erg Å. Inserting this relation to Equation 58 above yields

$$f_\lambda \left[\frac{\text{photons}}{\text{cm}^2\ \text{s Å}}\right] = 5.03 \times 10^7 \lambda[\text{Å}] F_\lambda \left[\frac{\text{erg}}{\text{cm}^2\ \text{s Å}}\right], \tag{59}$$

or

$$f_\lambda \left[\frac{\text{photons}}{\text{cm}^2\ \text{s nm}}\right] = 5.03 \times 10^8 \lambda[\text{Å}] F_\lambda \left[\frac{\text{erg}}{\text{cm}^2\ \text{s Å}}\right]. \tag{60}$$

B.2 From W m^{-2} nm^{-1} to photons cm^{-2} s^{-1} nm^{-1}

Within DPAC, fluxes in the Phoenix library are written in W m^{-2} nm^{-1}. To convert these units into photons cm^{-2} s^{-1} nm^{-1}, we simply only need to remember that $1\ \text{W} = 1\ \text{J s}^{-1}$, and that $1\ \text{m}^2 = 10^4\ \text{cm}^2$.

To relate energy with the number of photons, we use again Equation 55. We thus have

$$f_\lambda \left[\frac{\text{photons}}{\text{cm}^2\ \text{s}\ \text{nm}}\right] = 10^{-4}\frac{\lambda[\text{nm}]}{hc}F_\lambda \left[\frac{\text{Watt}}{\text{m}^2\ \text{nm}}\right]. \tag{61}$$

Here we can take again the value of hc, which is $hc = 1.986 \times 10^{-16}$ J nm. Equation 61 thus becomes

$$f_\lambda \left[\frac{\text{photons}}{\text{cm}^2\ \text{s}\ \text{nm}}\right] = 5.03 \times 10^{11}\lambda[\text{nm}]F_\lambda \left[\frac{\text{Watt}}{\text{m}^2\ \text{nm}}\right]. \tag{62}$$

B.3 From erg cm^{-2} s^{-1} Hz^{-1} to photons cm^{-2} s^{-1} nm^{-1}

Converting fluxes from erg cm^{-2} s^{-1} Hz^{-1} to photons cm^{-2} s^{-1} nm^{-1} is particularly useful for the Basel2 library since the fluxes are written in these units. To do the conversion, we first need to convert frequency into wavelength my means of

$$d\nu = \frac{c}{\lambda^2}d\lambda, \tag{63}$$

where c is the velocity of light. We can use this to transform erg cm^{-2} s^{-1} Hz^{-1} into erg cm^{-2} s^{-1} nm^{-1}:

$$F_\lambda \left[\frac{\text{erg}}{\text{cm}^2\ \text{s}\ \text{nm}}\right] = \frac{c}{\lambda^2}F_\nu \left[\frac{\text{erg}}{\text{cm}^2\ \text{s}\ \text{Hz}}\right]. \tag{64}$$

Here c is nm s^{-1}, i.e. $c = 3 \times 10^{17}$ nm s^{-1}, and λ in nanometer. To convert further into photons cm^{-2} s^{-1} nm^{-1}, we can use again the relation in Equation 55 to obtain

$$f_\lambda \left[\frac{\text{photons}}{\text{cm}^2\ \text{s}\ \text{nm}}\right] = \frac{1}{h\lambda}F_\nu \left[\frac{\text{erg}}{\text{cm}^2\ \text{s}\ \text{Hz}}\right]. \tag{65}$$

Here the value of h is in cgs, i.e. $h = 6.626 \times 10^{-27}$ erg s, and λ is in nanometer. Thus we have

$$f_\lambda \left[\frac{\text{photons}}{\text{cm}^2\ \text{s}\ \text{nm}}\right] = \frac{1.509 \times 10^{26}}{\lambda[\text{nm}]}F_\nu \left[\frac{\text{erg}}{\text{cm}^2\ \text{s}\ \text{Hz}}\right]. \tag{66}$$

C Examples of Ulysses BP/RP Spectra

All of the BP/RP spectra shown in this Appendix are noise-free and oversampled by a factor 8. Several types of stars are shown here. Stars with varying T_{eff} and $\log g$ (Figures 33–40), stars with varying $\log g$ and all other APs are kept the same (Figure 41), stars with varying [Fe/H] (Figure 42), and a star suffering various values of interstellar reddening (Figures 43–58). The input spectra are taken either from the Basel2 or the Phoenix spectral library.

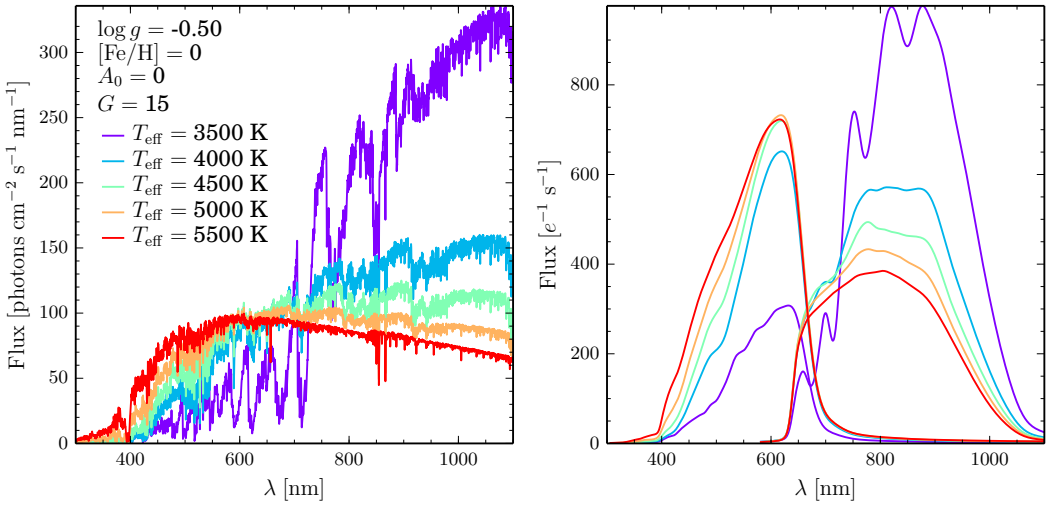


Figure 33: The noise-free BP/RP spectra (right plot) as simulated by Ulysses, for stars with various $T_{\rm eff}$ as indicated by the color coding. All other astrophysical parameters are kept the same: $[{\rm Fe/H}] = 0$, $A_0 = 0$, $G = 15$, and $\log g = -0.5$. The left plot shows the original spectra before it is observed by *Gaia*. Input spectra are all taken from the Phoenix library.

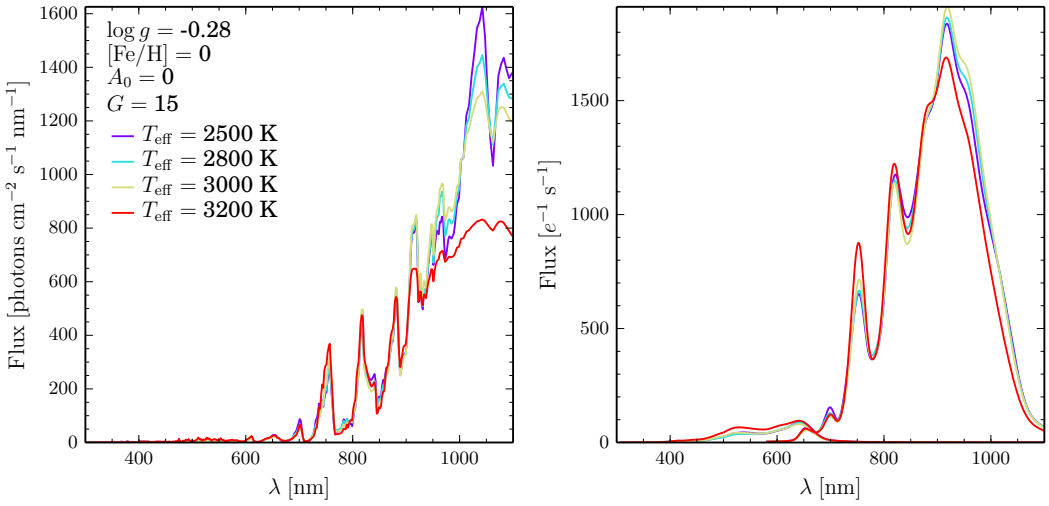


Figure 34: The same as in Figure 33, but for stars with $\log g = -0.28$ and various $T_{\rm eff}$ as indicated by the color coding. Input spectra are all taken from the Basel2 library.

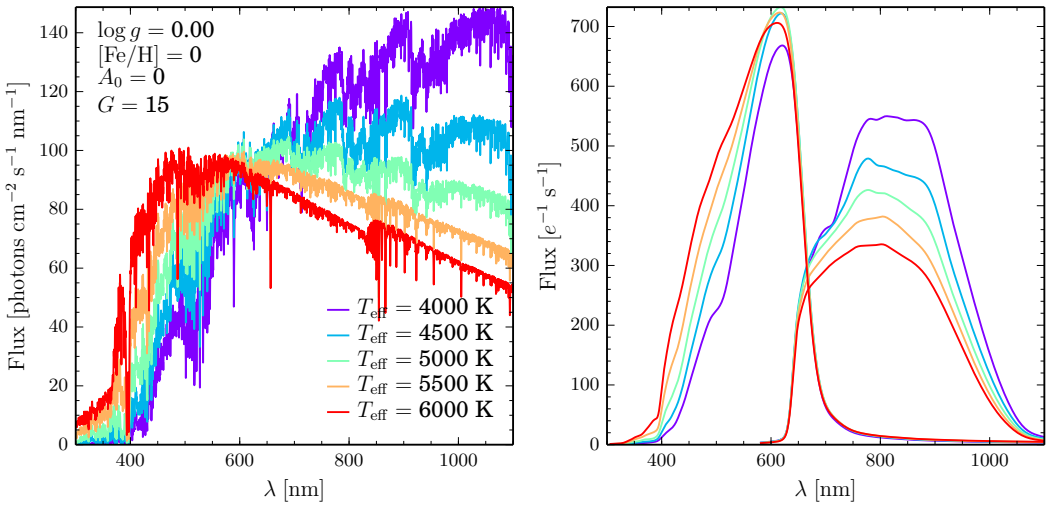


Figure 35: The same as in Figure 33, but for stars with $\log g = 0$ and various $T_{\rm eff}$ as indicated by the color coding. Input spectra are all taken from the Phoenix library.

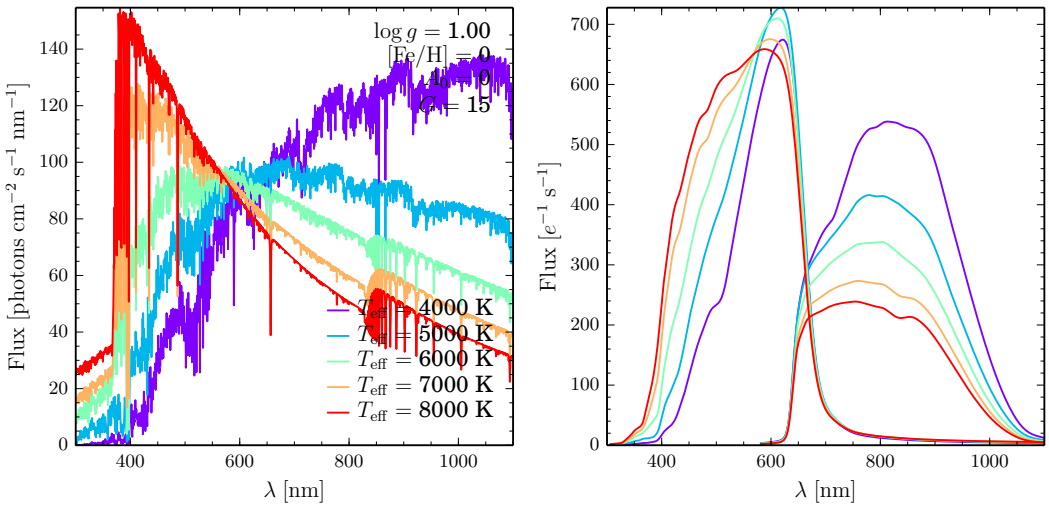


Figure 36: The same as in Figure 33, but for stars with $\log g = 1$ and various $T_{\rm eff}$ as indicated by the color coding. Input spectra are all taken from the Phoenix library.

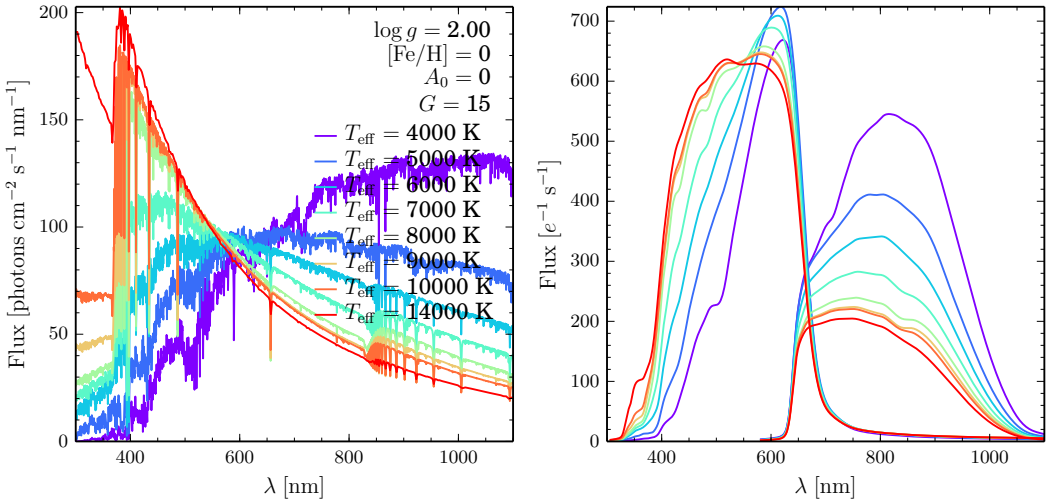


Figure 37: The same as in Figure 33, but for stars with $\log g = 2$ and various $T_{\rm eff}$ as indicated by the color coding. Input spectra for $T_{\rm eff} = 14\,000$ K are taken from the Basel2 library, the rest are from the Phoenix library.

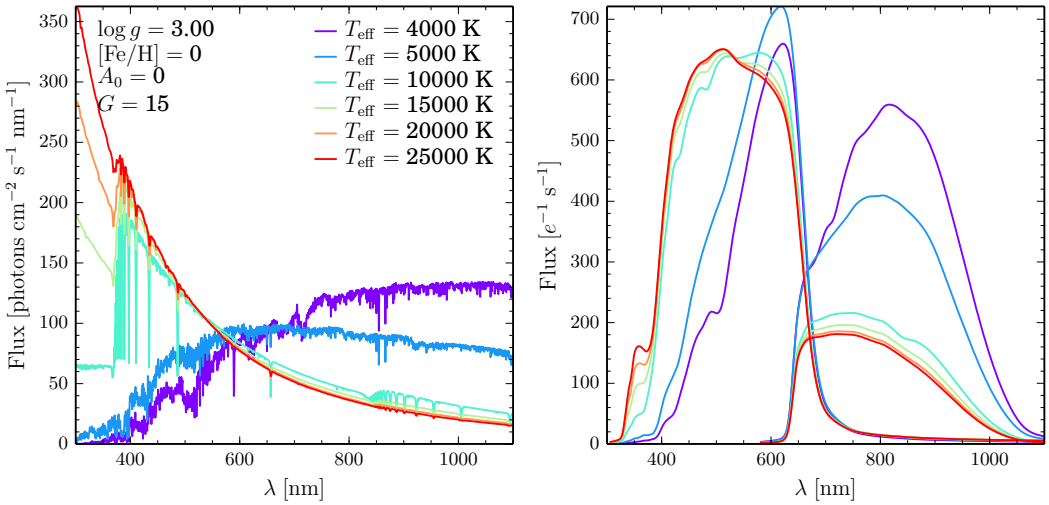


Figure 38: The same as in Figure 33, but for stars with $\log g = 3$ and various $T_{\rm eff}$ as indicated by the color coding. Input spectra for stars with $T_{\rm eff} \leq 10\,000$ K are taken from the Phoenix library, while the rest are taken from the Basel2 library.

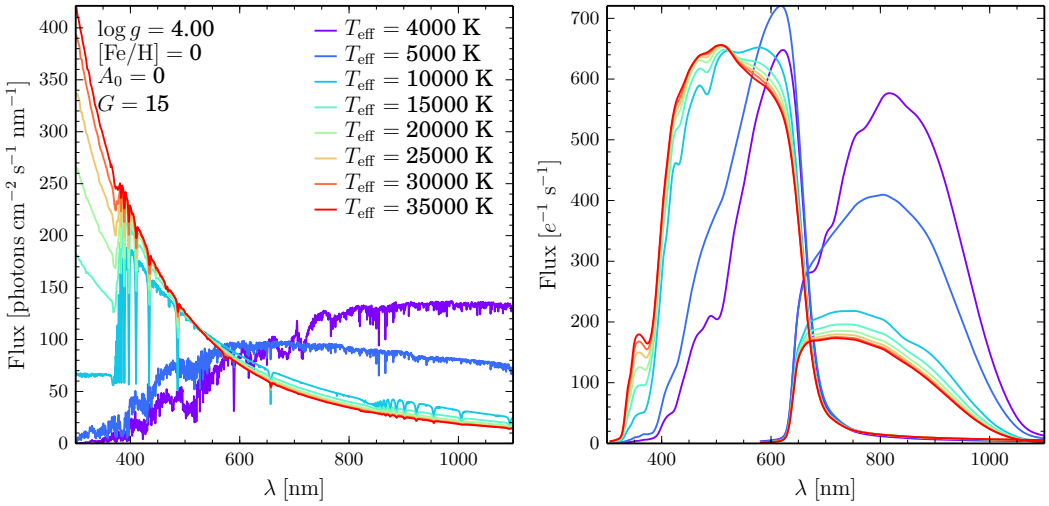


Figure 39: The same as in Figure 33, but for stars with $\log g = 4$ and various $T_{\rm eff}$ as indicated by the color coding. Input spectra for stars with $T_{\rm eff} \leq 10\,000$ K are taken from the Phoenix library, while the rest are taken from the Basel2 library.

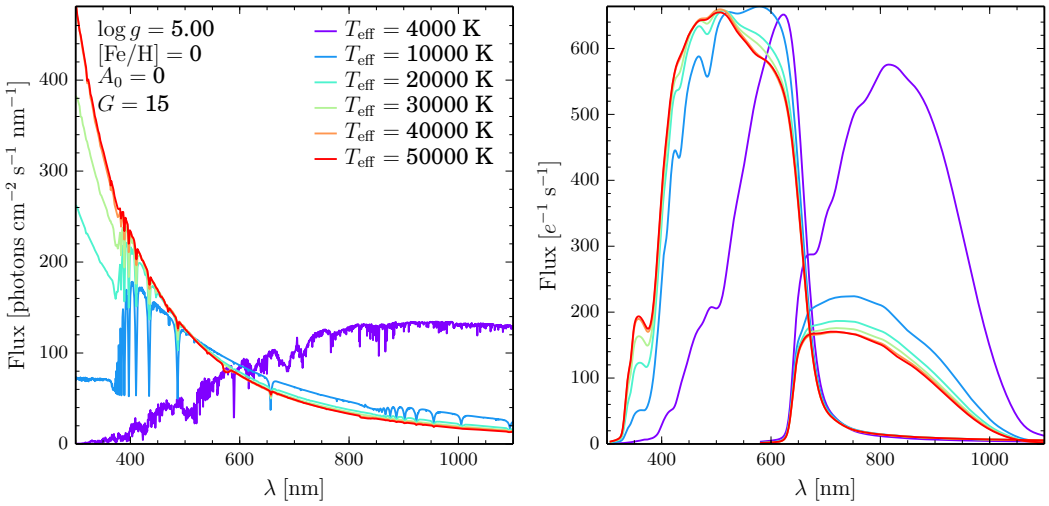


Figure 40: The same as in Figure 33, but for stars with $\log g = 5$ and various $T_{\rm eff}$ as indicated by the color coding. Input spectra for stars with $T_{\rm eff} \leq 10\,000$ K are taken from the Phoenix library, while the rest are taken from the Basel2 library.

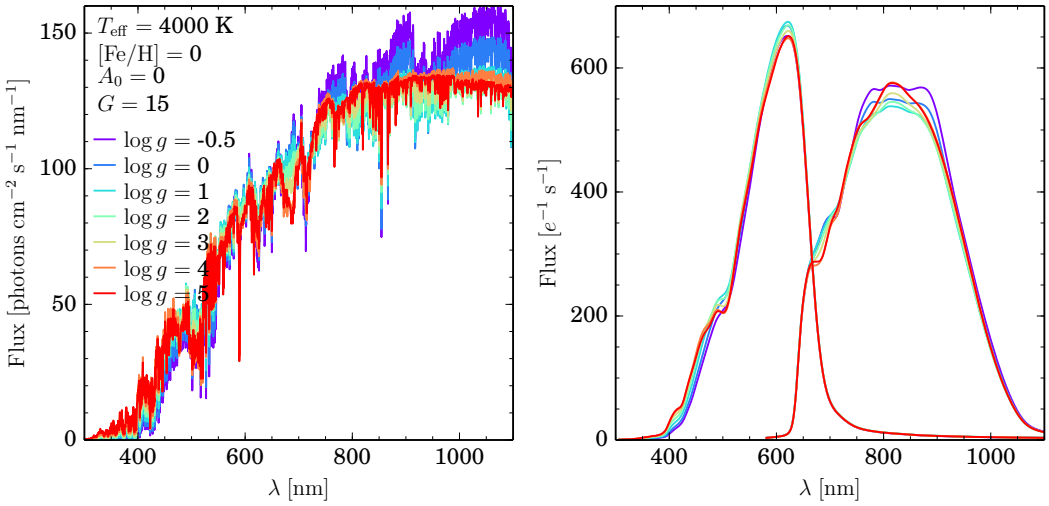


Figure 41: The noise-free BP/RP spectra (right plot) as simulated by Ulysses, for stars with various $\log g$ as indicated by the color coding. All other astrophysical parameters are kept the same: $T_{\rm eff} = 4000$ K, $[{\rm Fe/H}] = 0$, $A_0 = 0$, and $G = 15$. The left plot shows the original spectra before it is observed by *Gaia*. All input spectra are taken from the Phoenix library.

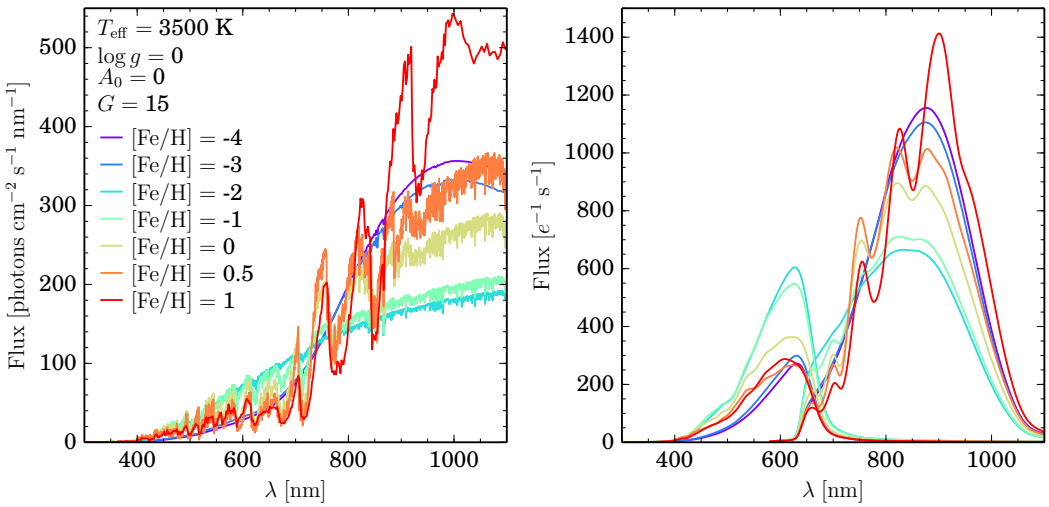


Figure 42: The noise-free BP/RP spectra (right plot) as simulated by Ulysses, for stars with various $[{\rm Fe/H}]$ as indicated by the color coding. All other astrophysical parameters are kept the same: $T_{\rm eff} = 3500$ K, $\log g = 0$, $A_0 = 0$, and $G = 15$. The left plot shows the original spectra before it is observed by *Gaia*. For stars with $[{\rm Fe/H}] = -4, -3, +1$, the input spectra are taken from the Basel2 library. The rest are from the Phoenix library.

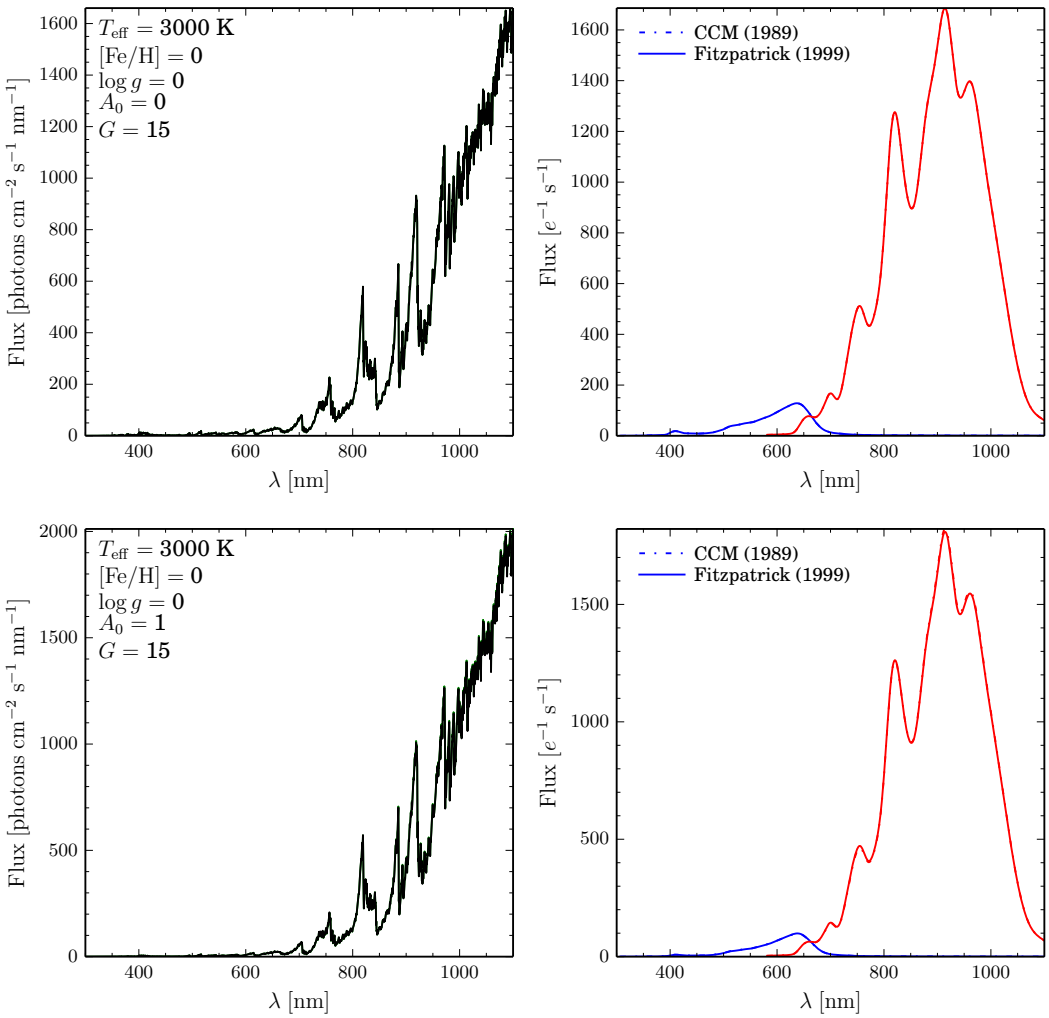


Figure 43: The noise-free BP/RP spectra (right plots, red and blue curves) of a star suffering various effects of reddening, i.e. $A_0 = \{0, 1, 3, 5, 10\}$. The reddened spectra before it is observed by *Gaia* are shown with the black curves on the left plot. The astrophysical parameters of the star are $T_{\rm eff} = 3000$ K, [Fe/H] $= 0$, $\log g = 0$, and $G = 15$. The top plots shows the star suffering no extinction, i.e. $A_0 = 0$, while the bottom shows the extinction of $A_0 = 1$.

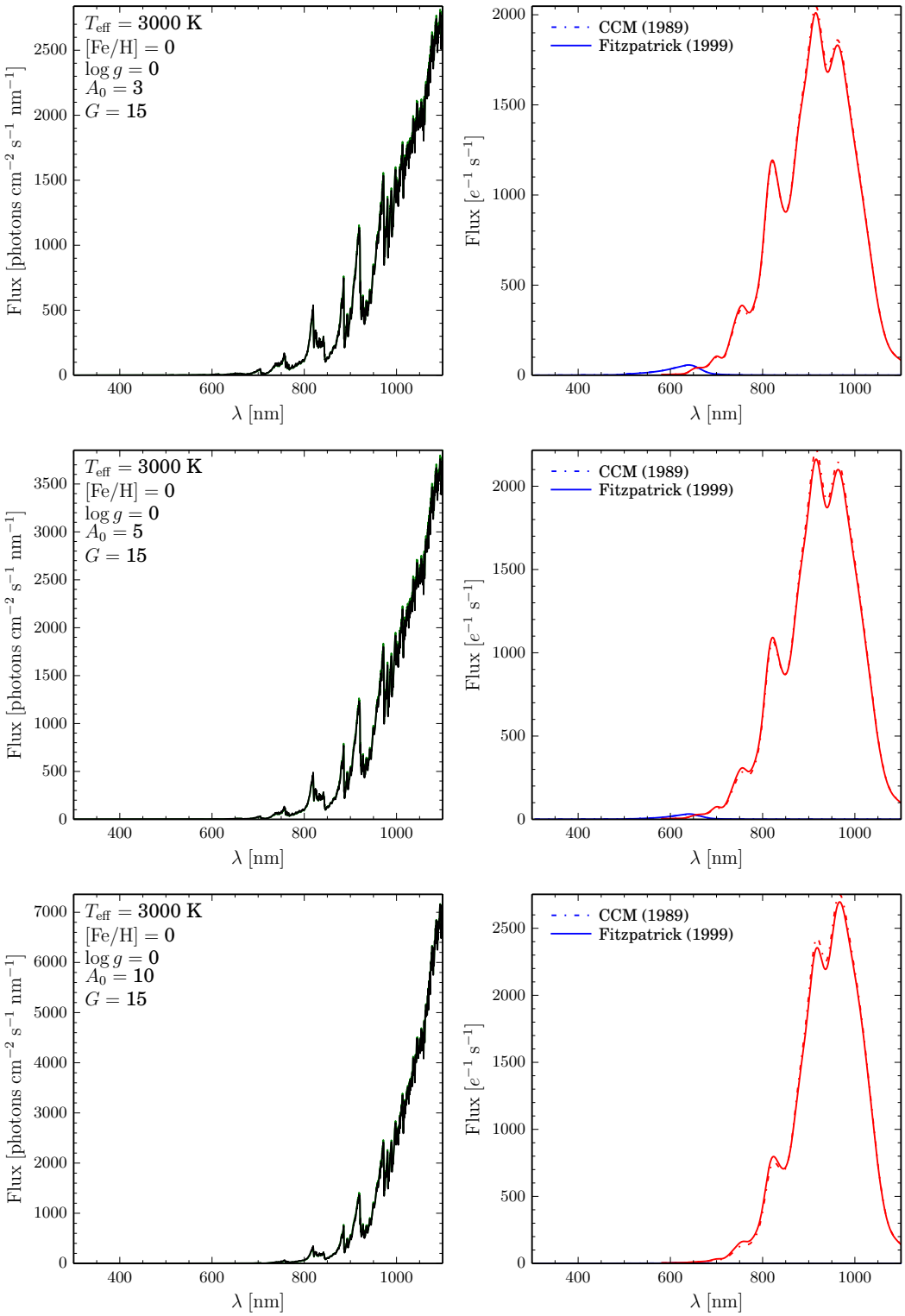


Figure 44: As in Figure 43, but for extinction $A_0 = 3$ (top plots), $A_0 = 5$ (middle plots). and $A_0 = 10$ (bottom plots).

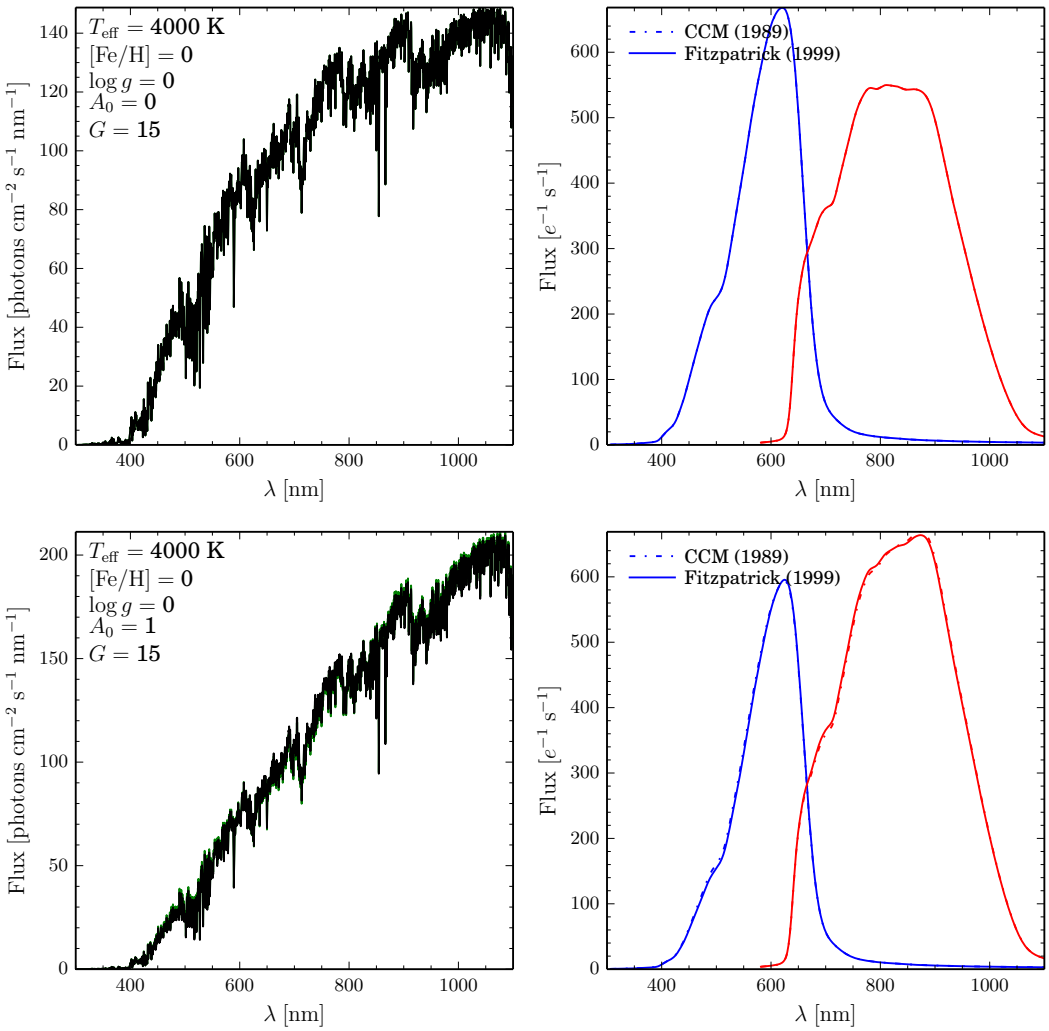


Figure 45: The noise-free BP/RP spectra (right plots, red and blue curves) of a star suffering various effects of reddening, i.e. $A_0 = \{0, 1, 3, 5, 10\}$. The reddened spectra before it is observed by *Gaia* are shown with the black curves on the left plot. The astrophysical parameters of the star are $T_{\rm eff} = 4000$ K, [Fe/H] $= 0$, $\log g = 0$, and $G = 15$. The top plots shows the star suffering no extinction, i.e. $A_0 = 0$, while the bottom shows the extinction of $A_0 = 1$.

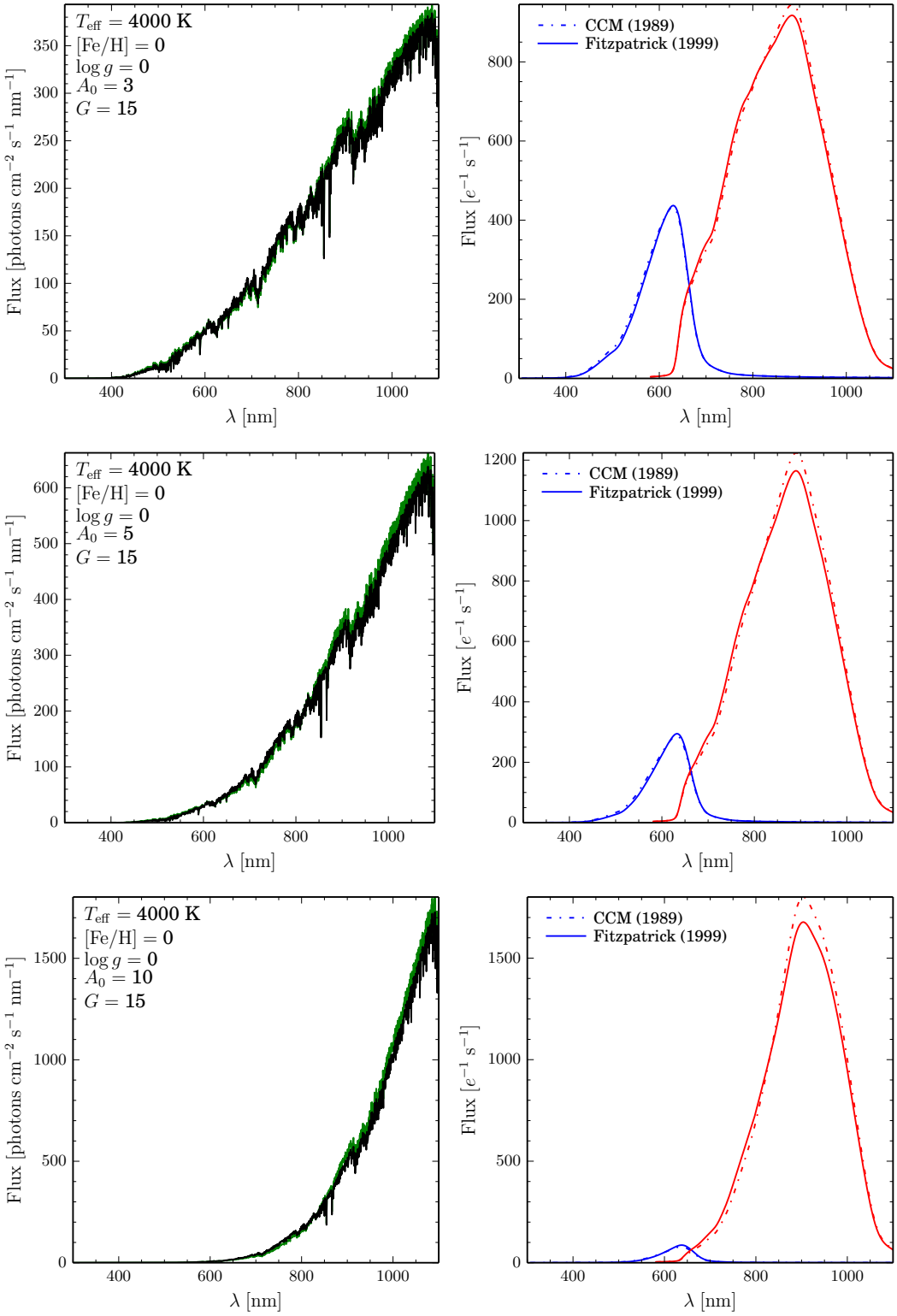


Figure 46: As in Figure 45, but for extinction $A_0 = 3$ (top plots), $A_0 = 5$ (middle plots). and $A_0 = 10$ (bottom plots).

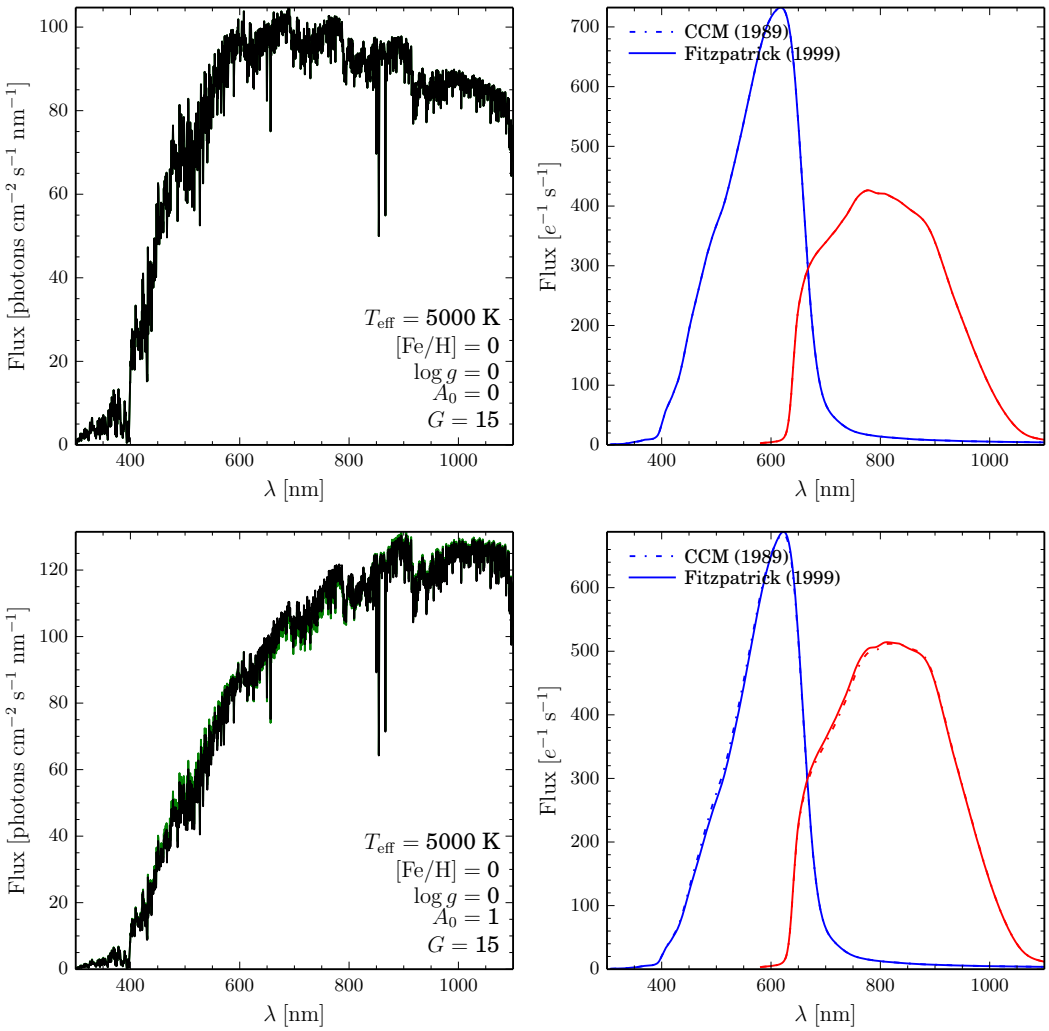


Figure 47: The noise-free BP/RP spectra (right plots, red and blue curves) of a star suffering various effects of reddening, i.e. $A_0 = \{0, 1, 3, 5, 10\}$. The reddened spectra before it is observed by *Gaia* are shown with the black curves on the left plot. The astrophysical parameters of the star are $T_{\rm eff} = 5000$ K, [Fe/H] $= 0$, $\log g = 0$, and $G = 15$. The top plots shows the star suffering no extinction, i.e. $A_0 = 0$, while the bottom shows the extinction of $A_0 = 1$.

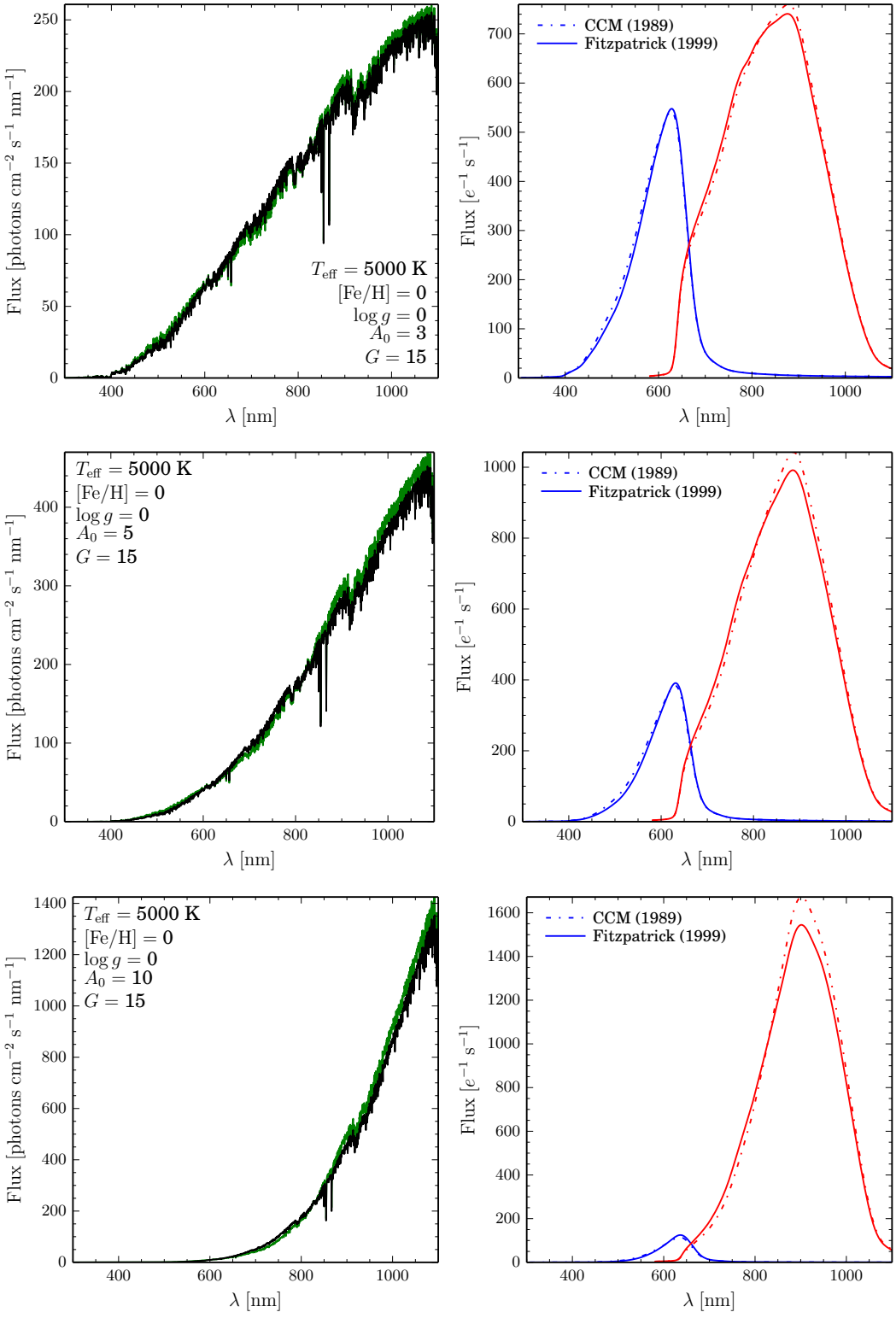


Figure 48: As in Figure 47, but for extinction $A_0 = 3$ (top plots), $A_0 = 5$ (middle plots). and $A_0 = 10$ (bottom plots).

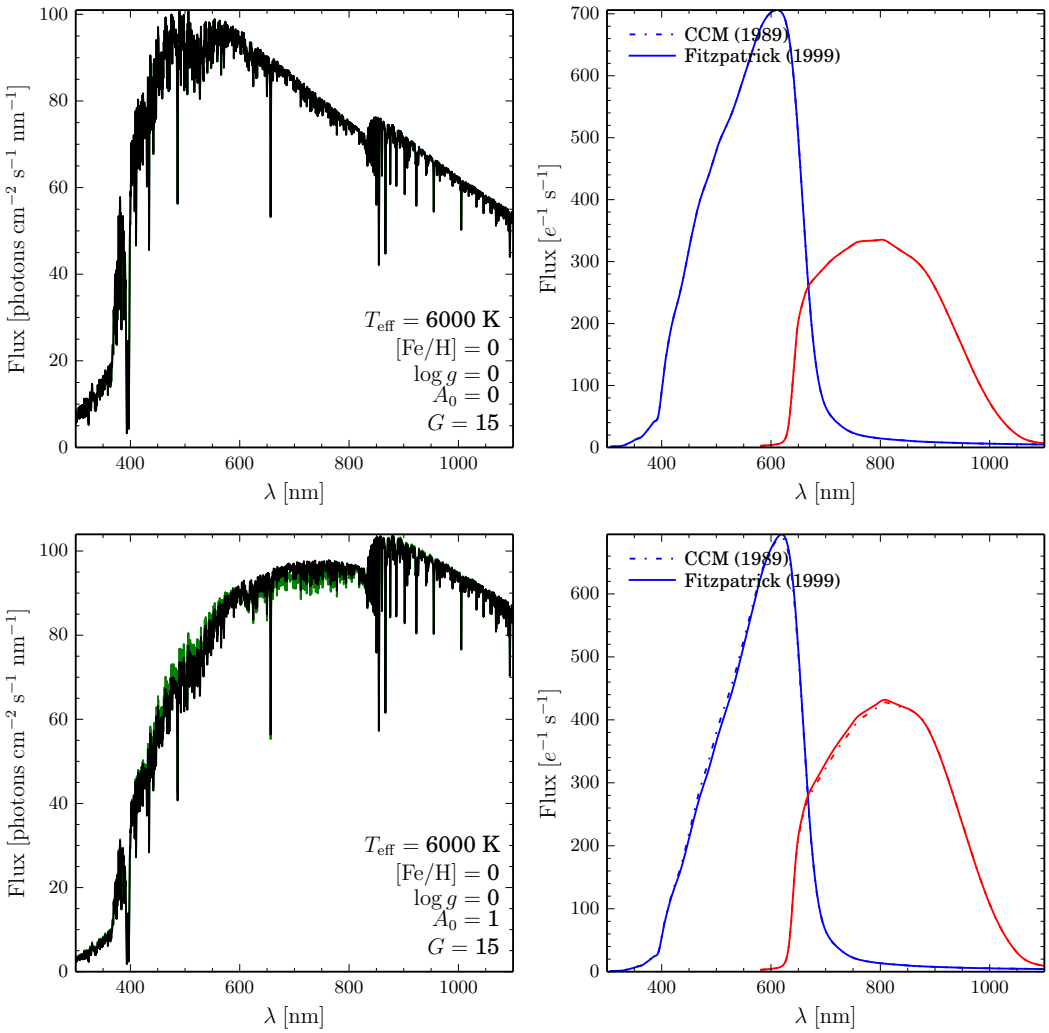


Figure 49: The noise-free BP/RP spectra (right plots, red and blue curves) of a star suffering various effects of reddening, i.e. $A_0 = \{0, 1, 3, 5, 10\}$. The reddened spectra before it is observed by *Gaia* are shown with the black curves on the left plot. The astrophysical parameters of the star are $T_{\rm eff} = 6000$ K, $[{\rm Fe/H}] = 0$, $\log g = 0$, and $G = 15$. The top plots shows the star suffering no extinction, i.e. $A_0 = 0$, while the bottom shows the extinction of $A_0 = 1$.

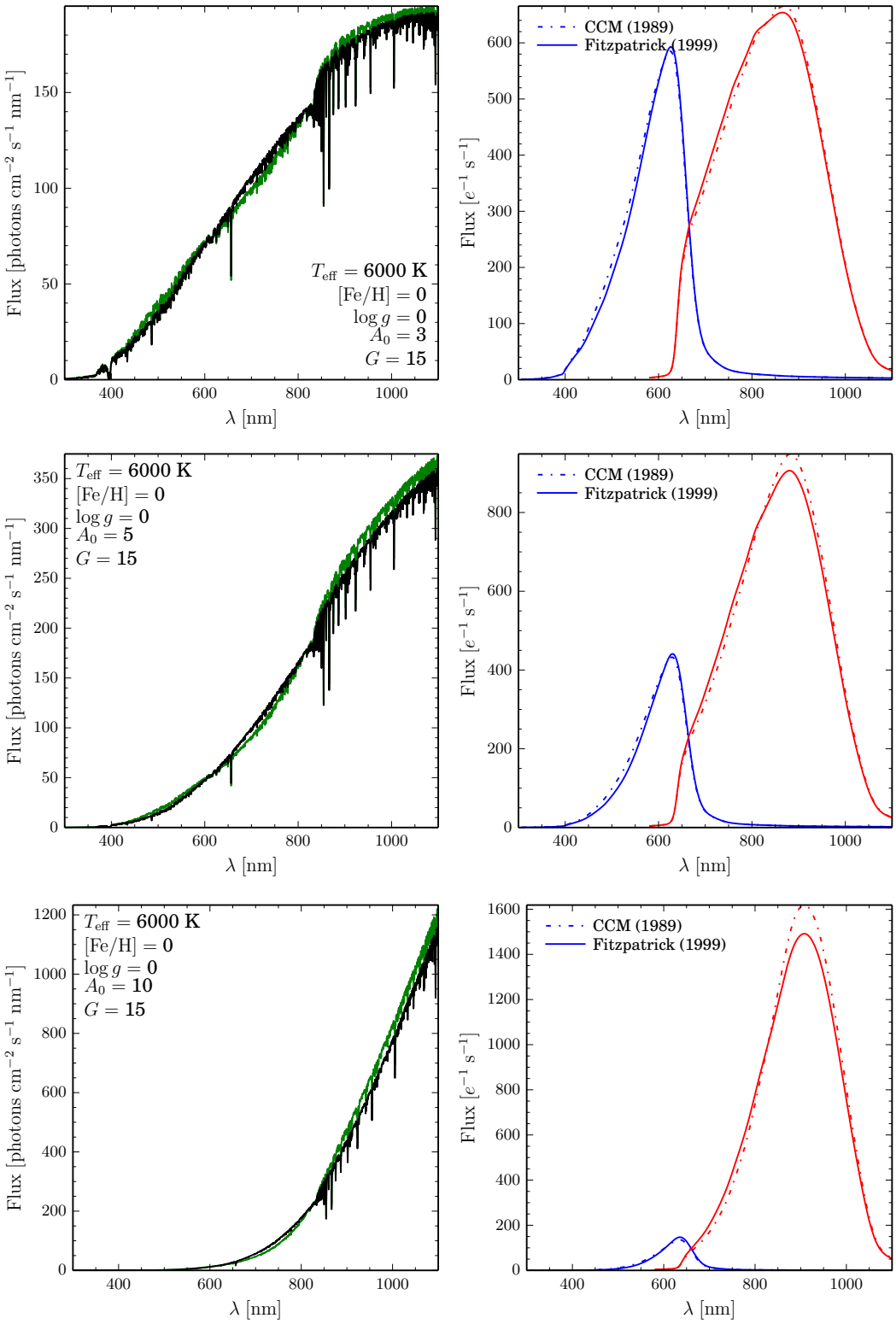


Figure 50: As in Figure 49, but for extinction $A_0 = 3$ (top plots), $A_0 = 5$ (middle plots). and $A_0 = 10$ (bottom plots).

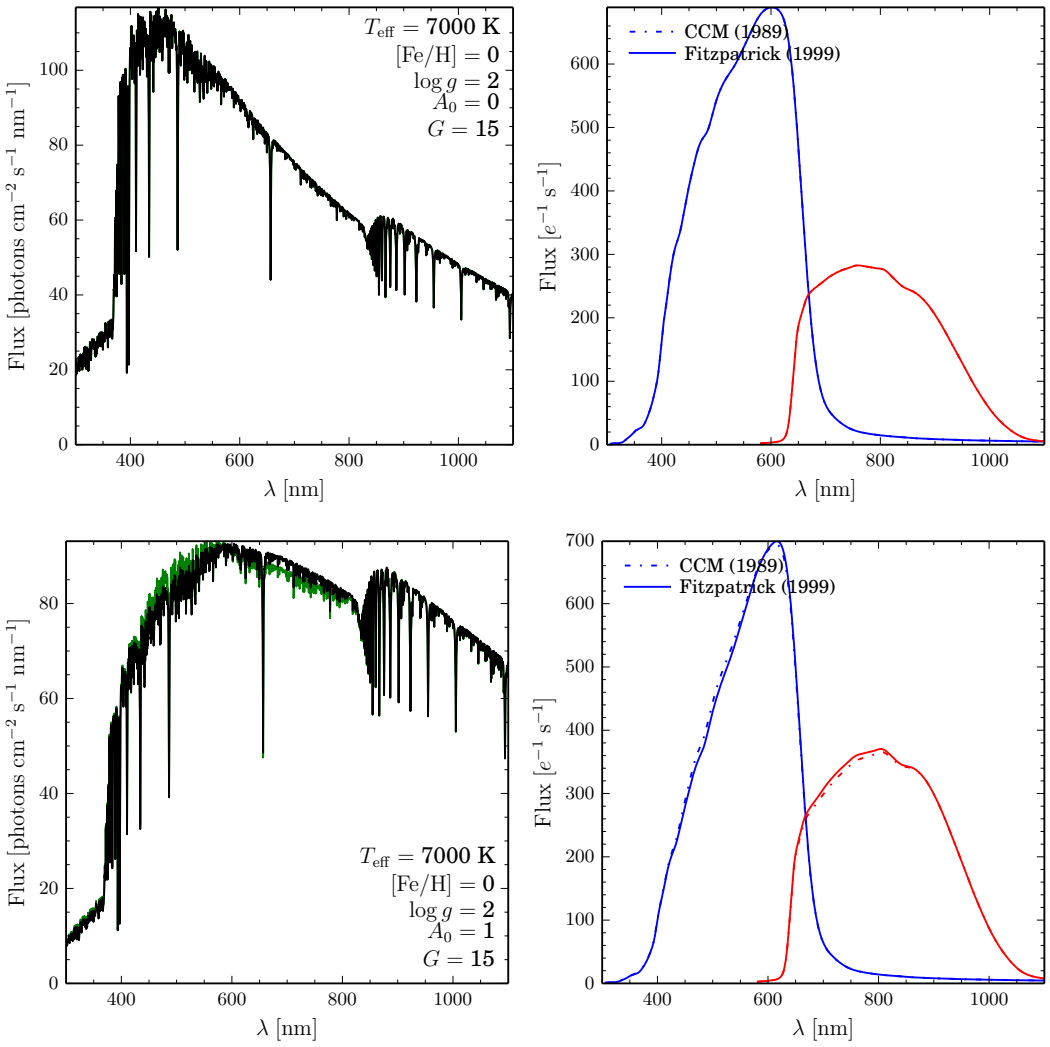


Figure 51: The noise-free BP/RP spectra (right plots, red and blue curves) of a star suffering various effects of reddening, i.e. $A_0 = \{0, 1, 3, 5, 10\}$. The reddened spectra before it is observed by *Gaia* are shown with the black curves on the left plot. The astrophysical parameters of the star are $T_{\rm eff} = 7000$ K, $[{\rm Fe/H}] = 0$, $\log g = 0$, and $G = 15$. The top plots shows the star suffering no extinction, i.e. $A_0 = 0$, while the bottom shows the extinction of $A_0 = 1$.

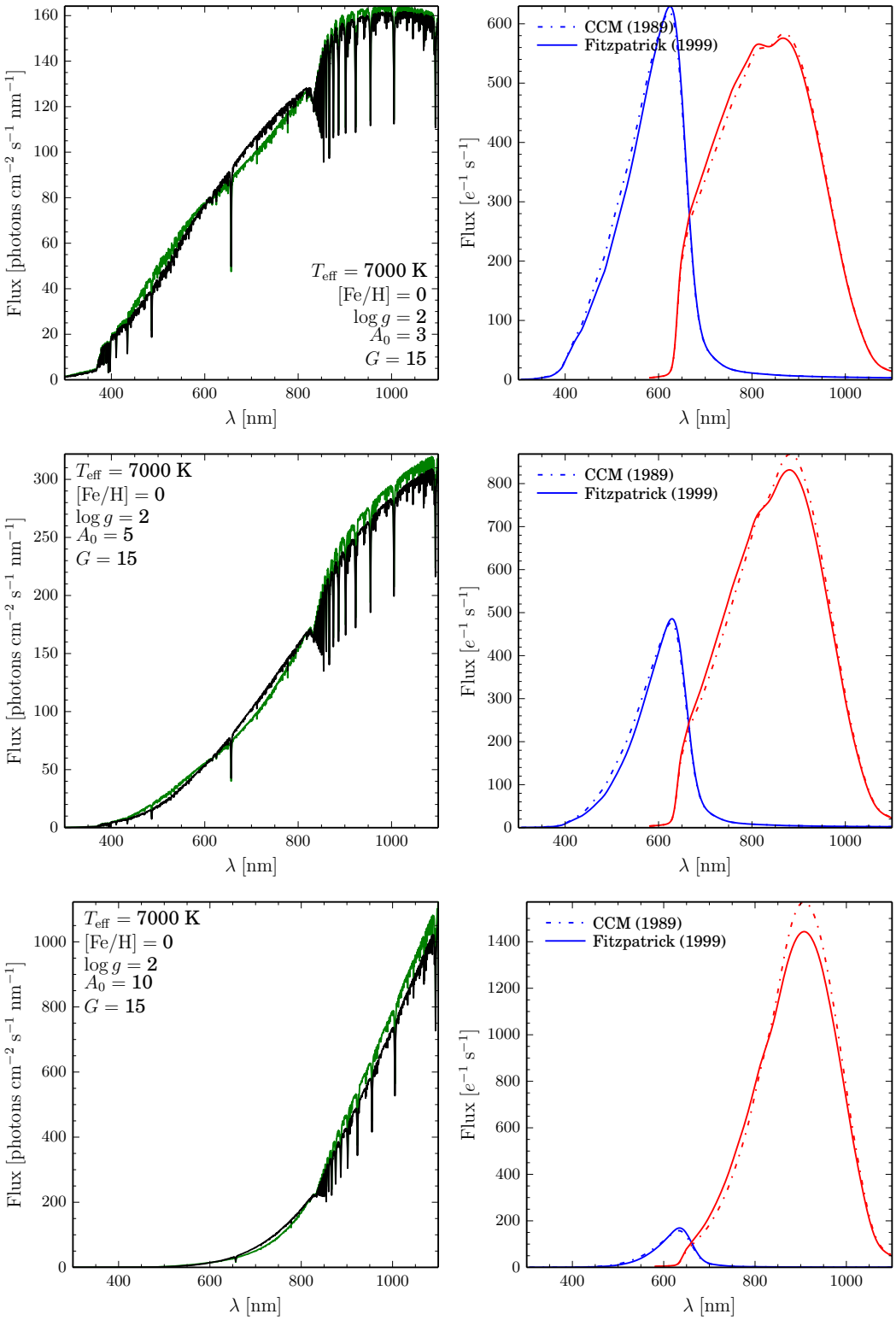


Figure 52: As in Figure 51, but for extinction $A_0 = 3$ (top plots), $A_0 = 5$ (middle plots). and $A_0 = 10$ (bottom plots).

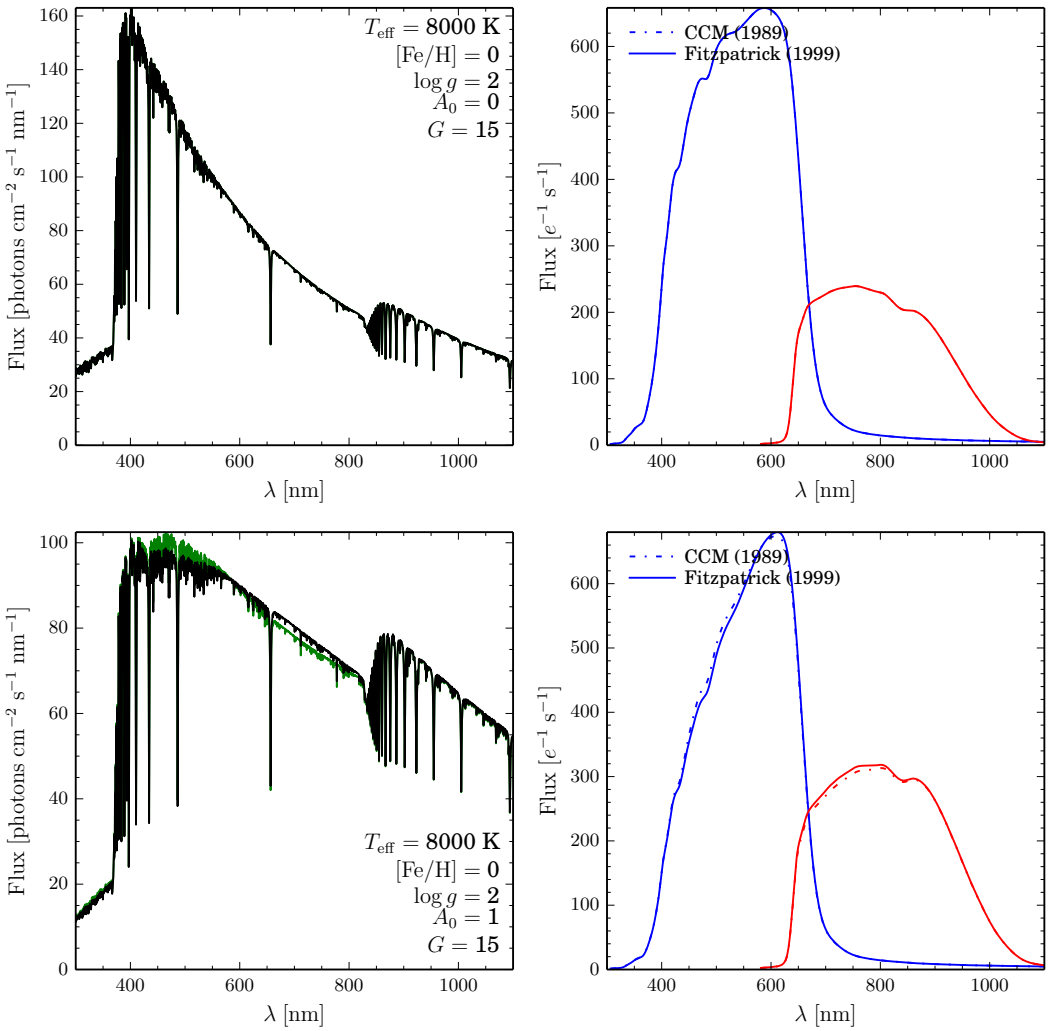


Figure 53: The noise-free BP/RP spectra (right plots, red and blue curves) of a star suffering various effects of reddening, i.e. $A_0 = \{0, 1, 3, 5, 10\}$. The reddened spectra before it is observed by *Gaia* are shown with the black curves on the left plot. The astrophysical parameters of the star are $T_{\rm eff} = 8000$ K, $[{\rm Fe/H}] = 0$, $\log g = 0$, and $G = 15$. The top plots shows the star suffering no extinction, i.e. $A_0 = 0$, while the bottom shows the extinction of $A_0 = 1$.

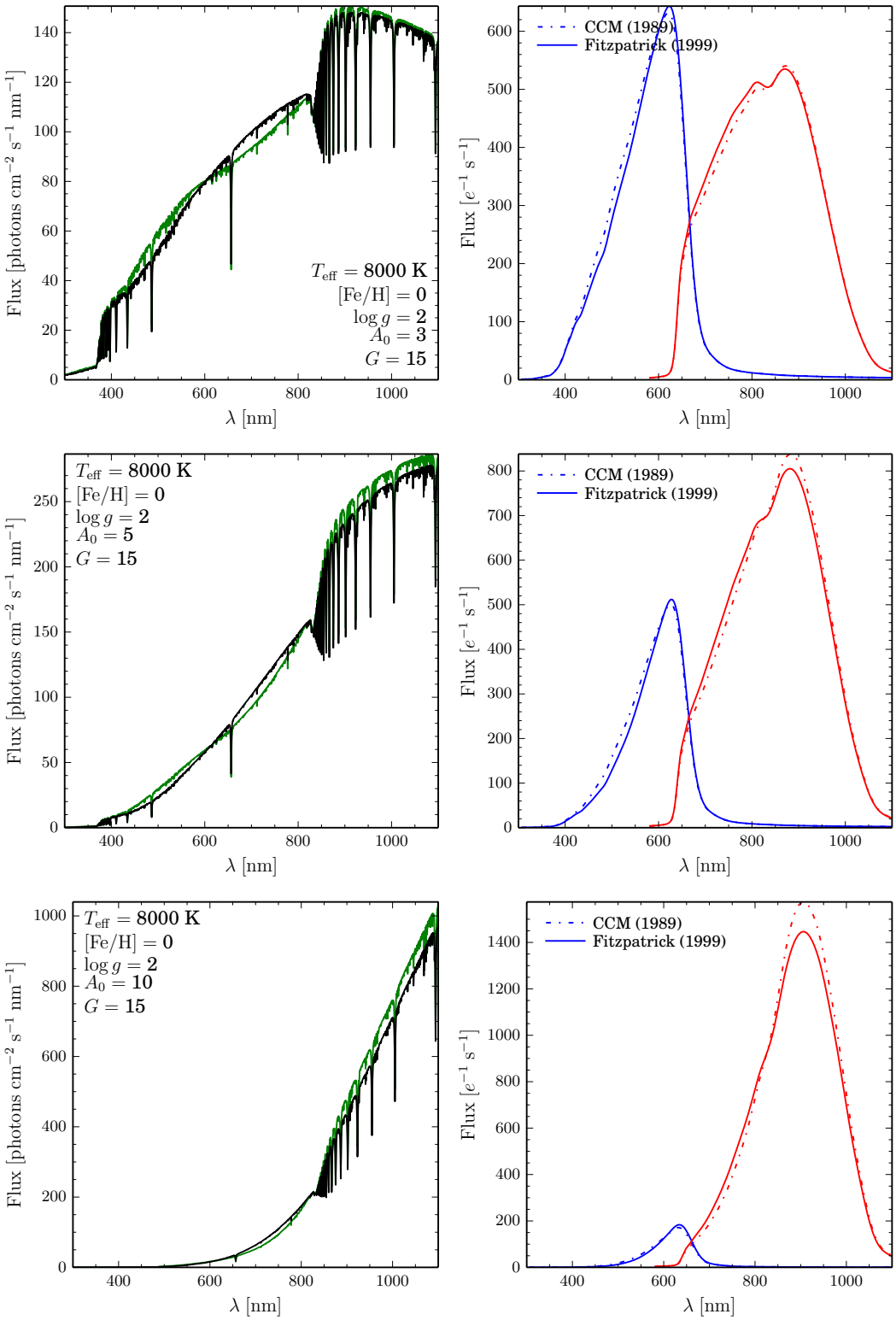


Figure 54: As in Figure 53, but for extinction $A_0 = 3$ (top plots), $A_0 = 5$ (middle plots). and $A_0 = 10$ (bottom plots).

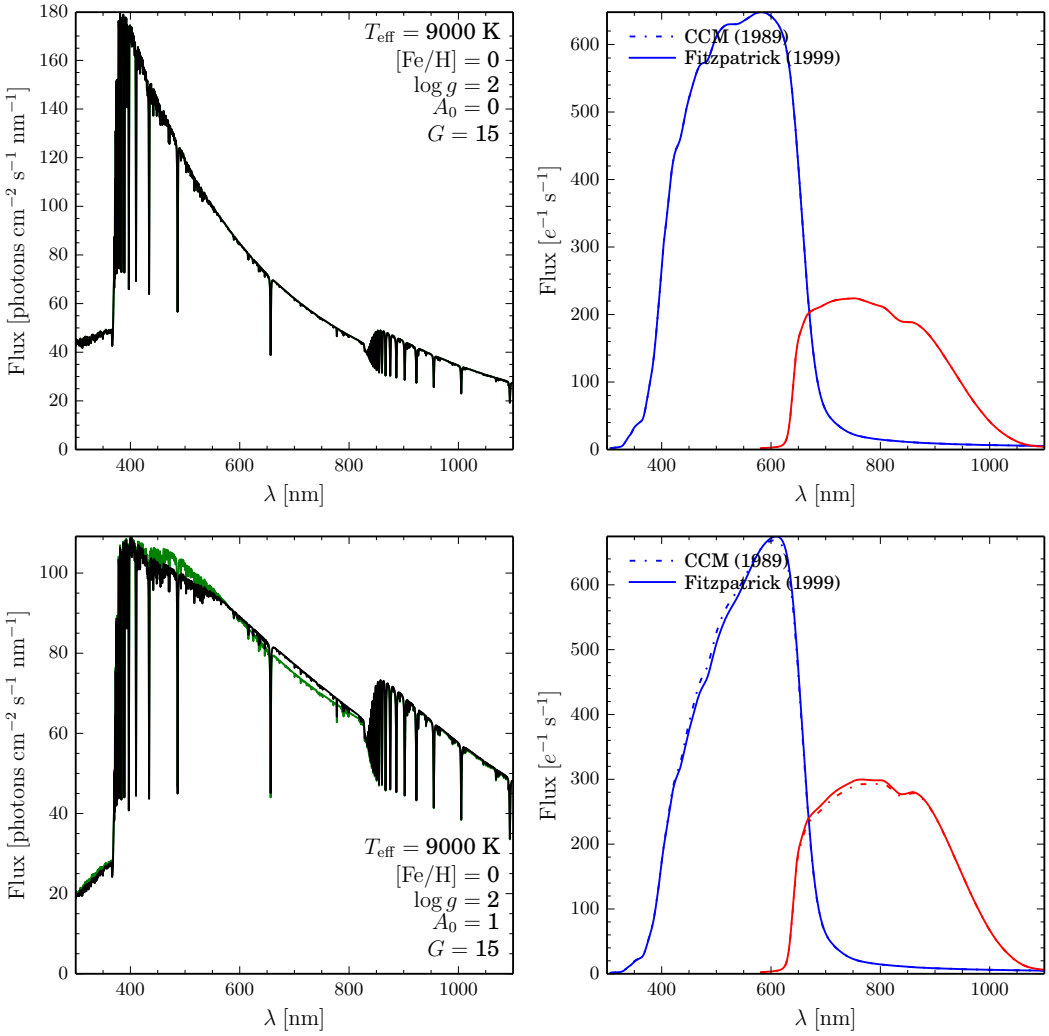


Figure 55: The noise-free BP/RP spectra (right plots, red and blue curves) of a star suffering various effects of reddening, i.e. $A_0 = \{0, 1, 3, 5, 10\}$. The reddened spectra before it is observed by *Gaia* are shown with the black curves on the left plot. The astrophysical parameters of the star are $T_{\rm eff} = 9000$ K, $[{\rm Fe/H}] = 0$, $\log g = 0$, and $G = 15$. The top plots shows the star suffering no extinction, i.e. $A_0 = 0$, while the bottom shows the extinction of $A_0 = 1$.

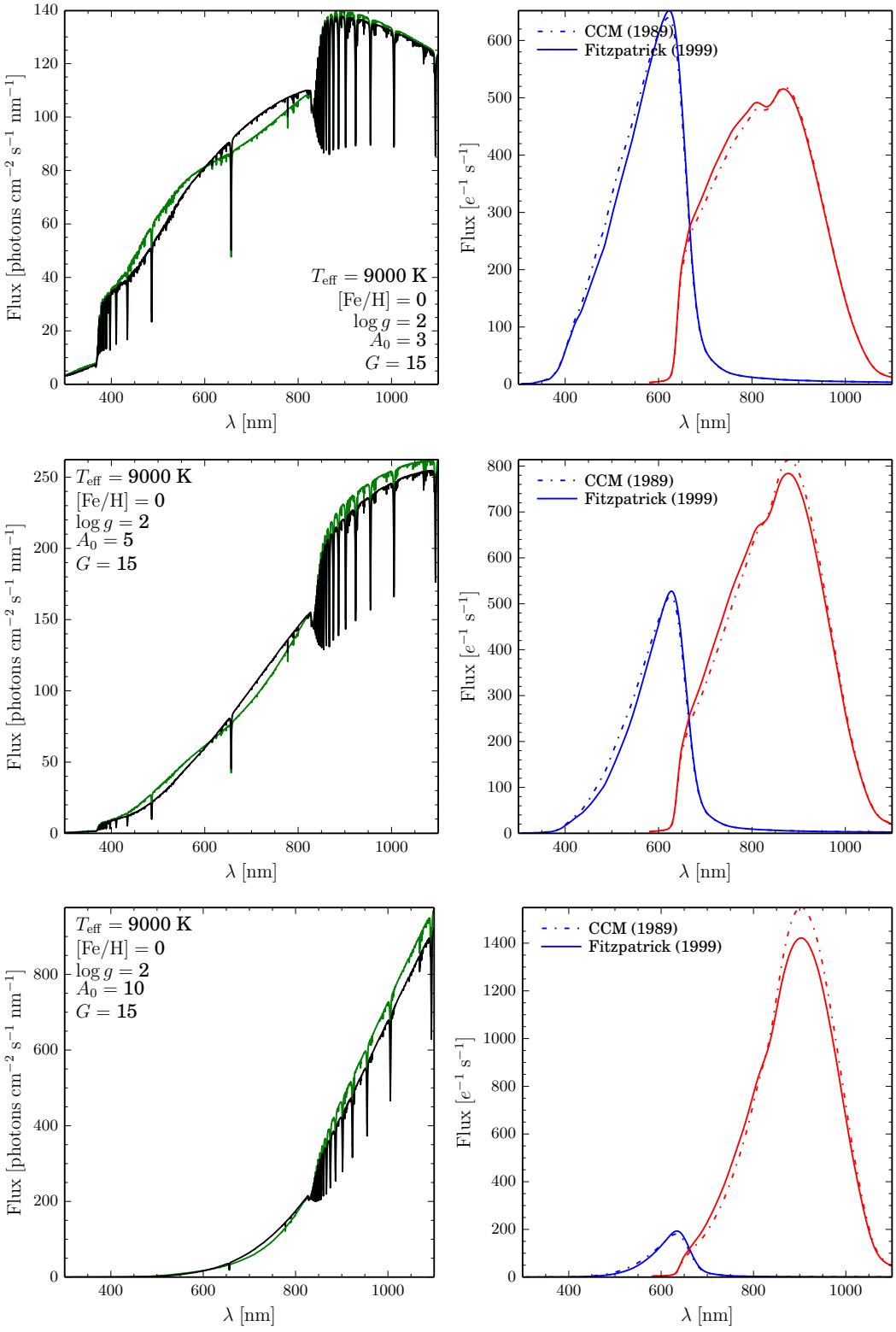


Figure 56: As in Figure 55, but for extinction $A_0 = 3$ (top plots), $A_0 = 5$ (middle plots). and $A_0 = 10$ (bottom plots).

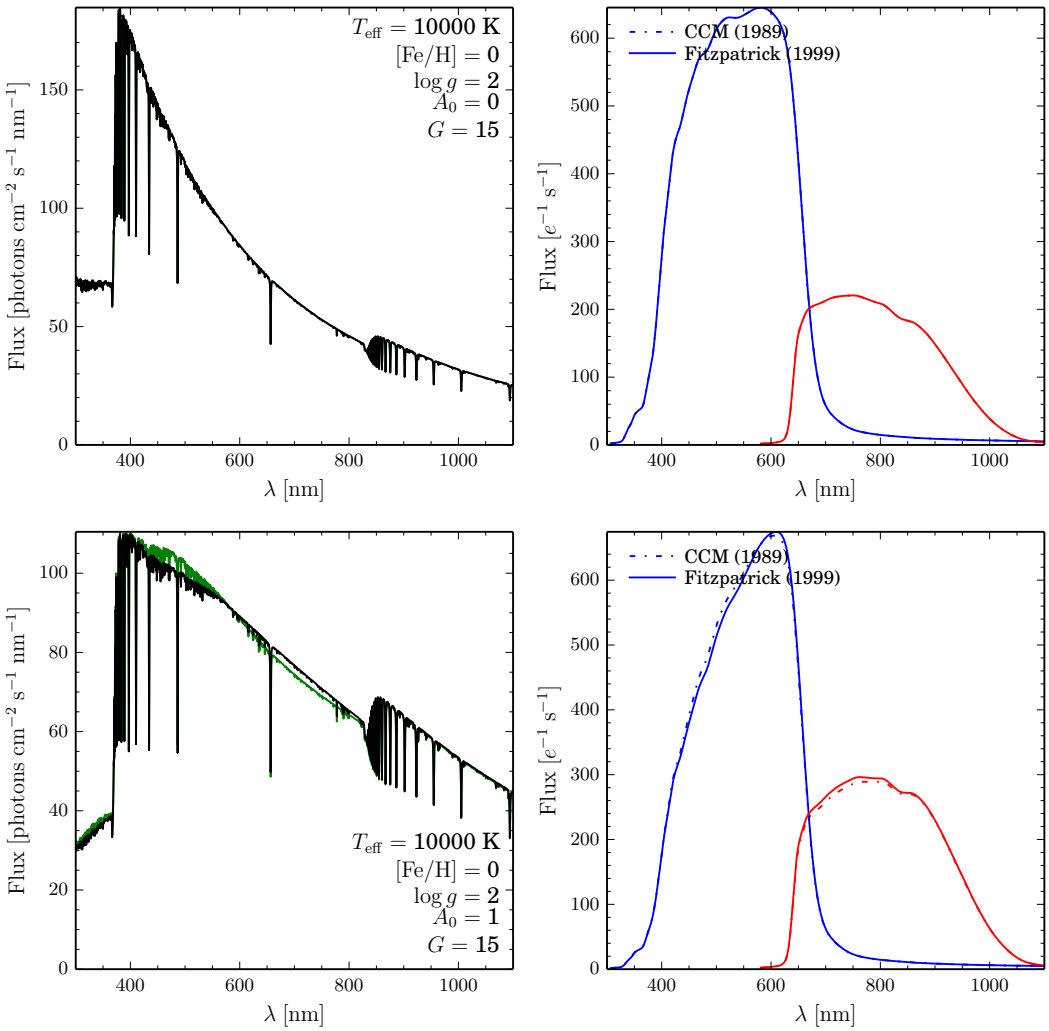


Figure 57: The noise-free BP/RP spectra (right plots, red and blue curves) of a star suffering various effects of reddening, i.e. $A_0 = \{0, 1, 3, 5, 10\}$. The reddened spectra before it is observed by *Gaia* are shown with the black curves on the left plot. The astrophysical parameters of the star are $T_{\rm eff} = 10000$ K, $[{\rm Fe/H}] = 0$, $\log g = 0$, and $G = 15$. The top plots shows the star suffering no extinction, i.e. $A_0 = 0$, while the bottom shows the extinction of $A_0 = 1$.

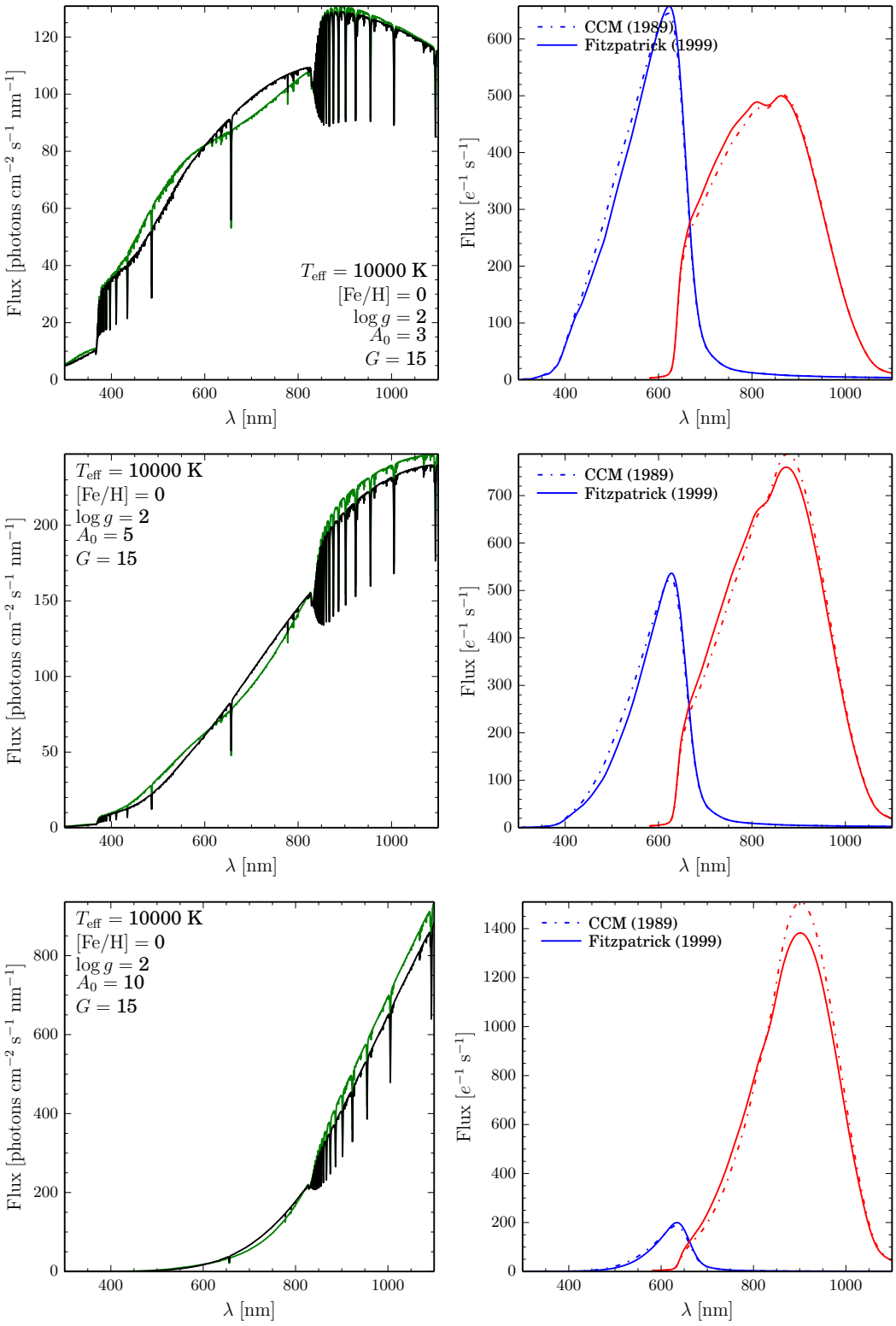


Figure 58: As in Figure 57, but for extinction $A_0 = 3$ (top plots), $A_0 = 5$ (middle plots). and $A_0 = 10$ (bottom plots).