

Automated parametrization of stellar spectra: algorithms in the Gaia context and their applicability

Alejandra Recio-Blanco

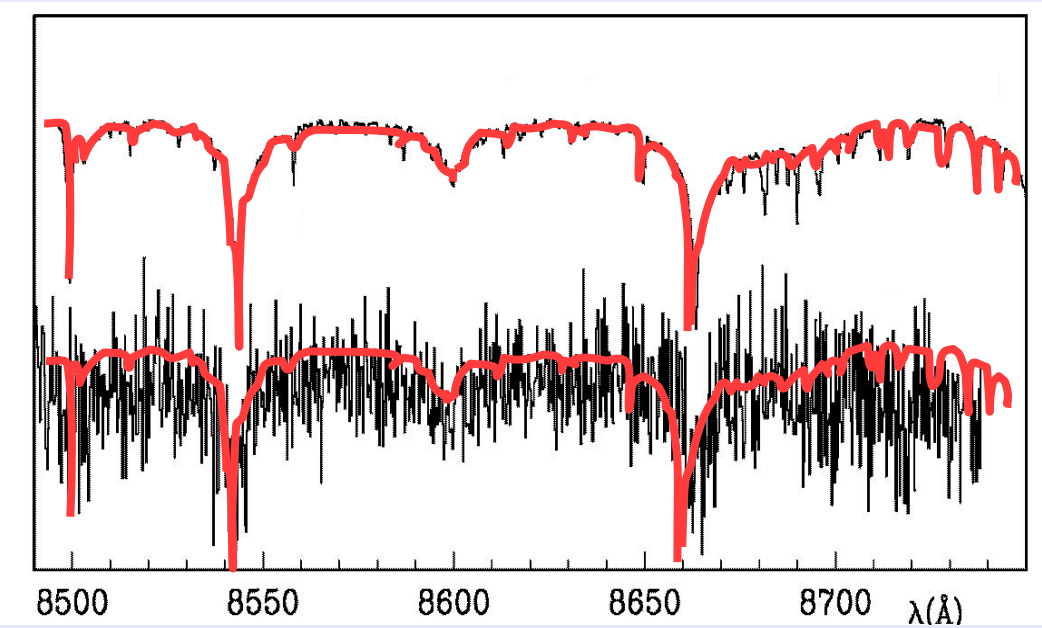
Collaborators: Patrick de Laverny, Albert Bijaoui

Observatoire de la Côte d'Azur

Automated stellar parametrization

1. What is it?

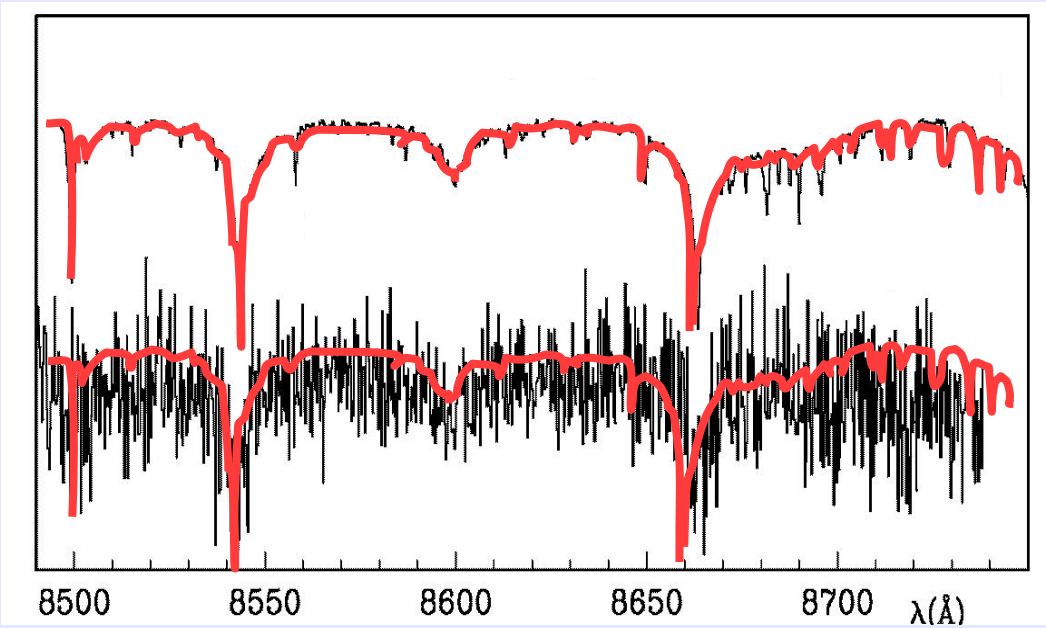
To determine the stellar parameters (effective temperature, surface gravity, global metallicity, individual chemical abundances) **that best fit an observed spectrum with a synthetic one.**



Automated stellar parametrization

1. What is it?

To determine the stellar parameters (effective temperature, surface gravity, global metallicity, individual chemical abundances) **that best fit an observed spectrum with a synthetic one.**



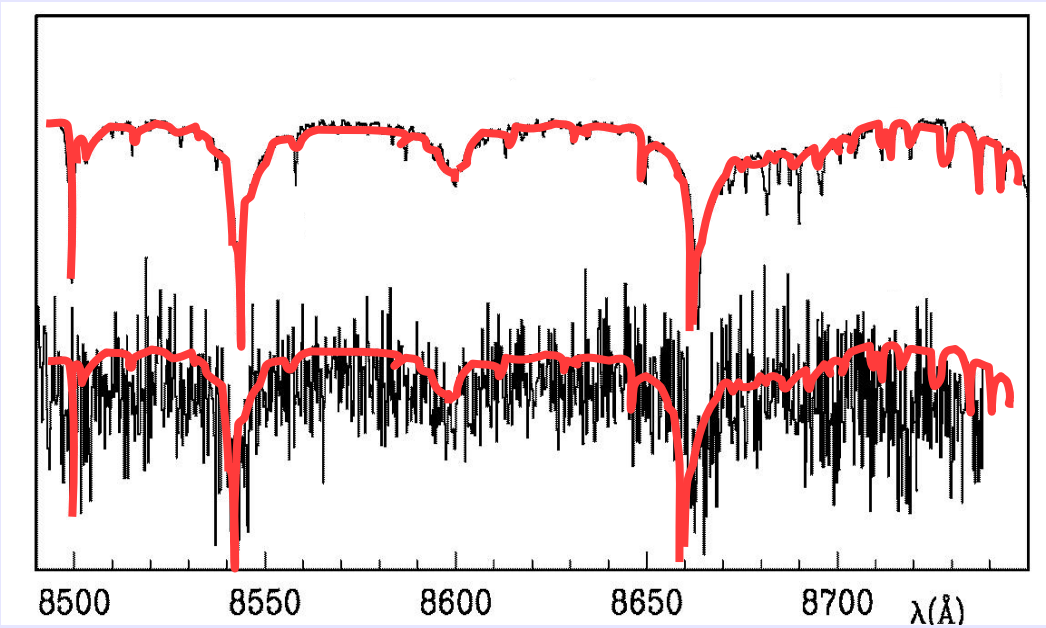
Distance minimization:

$$D(\Theta) = \sum_{l=1,L} [O(l) - S(l,\Theta)]^2.$$

Automated stellar parametrization

1. What is it?

To determine the stellar parameters (effective temperature, surface gravity, global metallicity, individual chemical abundances) **that best fit an observed spectrum with a synthetic one.**



Complex physics :

Model atmospheres

Atomic/molecular line lists

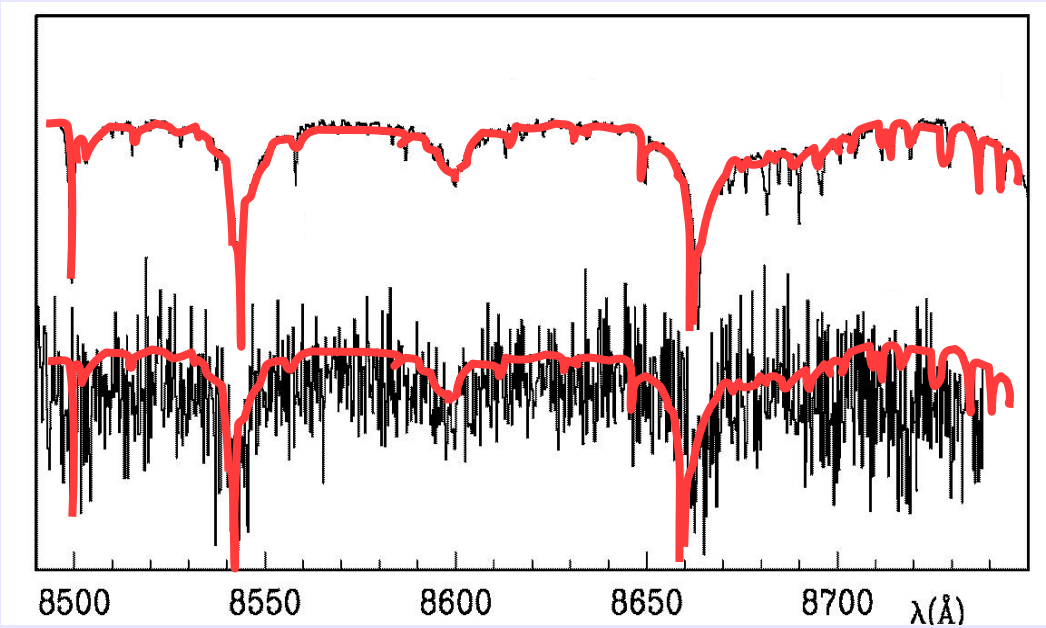
Atmospheric parameters: **Teff, log g, [M/H],...**

Impossible to work with analytical expressions !

Automated stellar parametrization

1. What is it?

To determine the stellar parameters (effective temperature, surface gravity, global metallicity, individual chemical abundances) **that best fit an observed spectrum with a synthetic one .**



Complex physics :

Model atmospheres

Atomic/molecular line lists

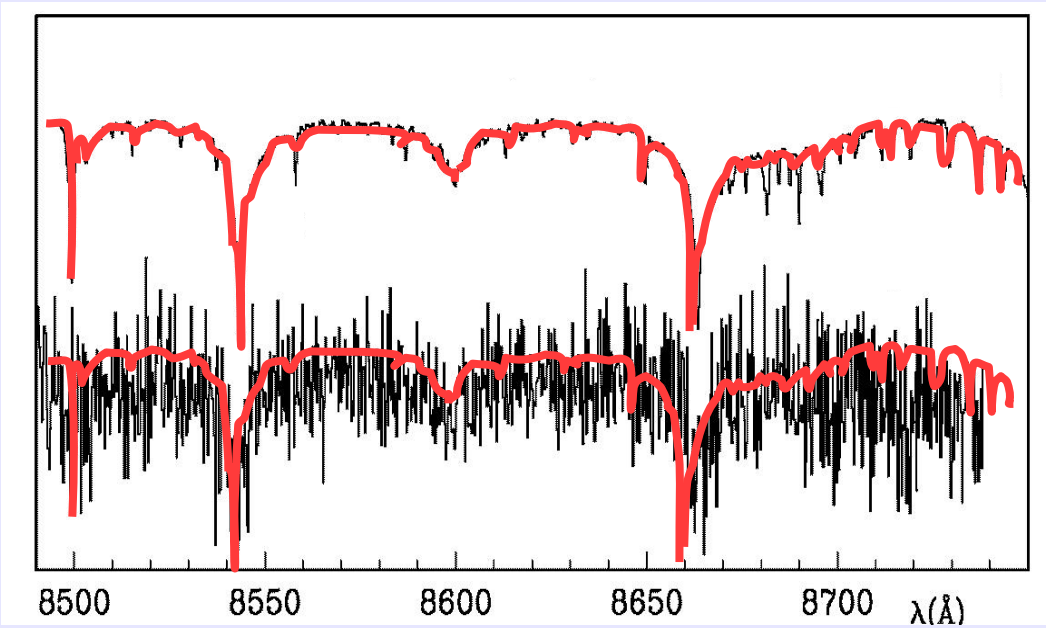
Atmospheric parameters: **Teff, log g, [M/H]**

Impossible to work with analytical expressions ! => **Grid of synthetic spectra**

Automated stellar parametrization

1. What is it?

To determine the stellar parameters (effective temperature, surface gravity, global metallicity, individual chemical abundances) **that best fit an observed spectrum with a synthetic one.**





Warning!

Any method of parametrization RELIES on the physics of the synthetic spectra grid, that has to be verified.

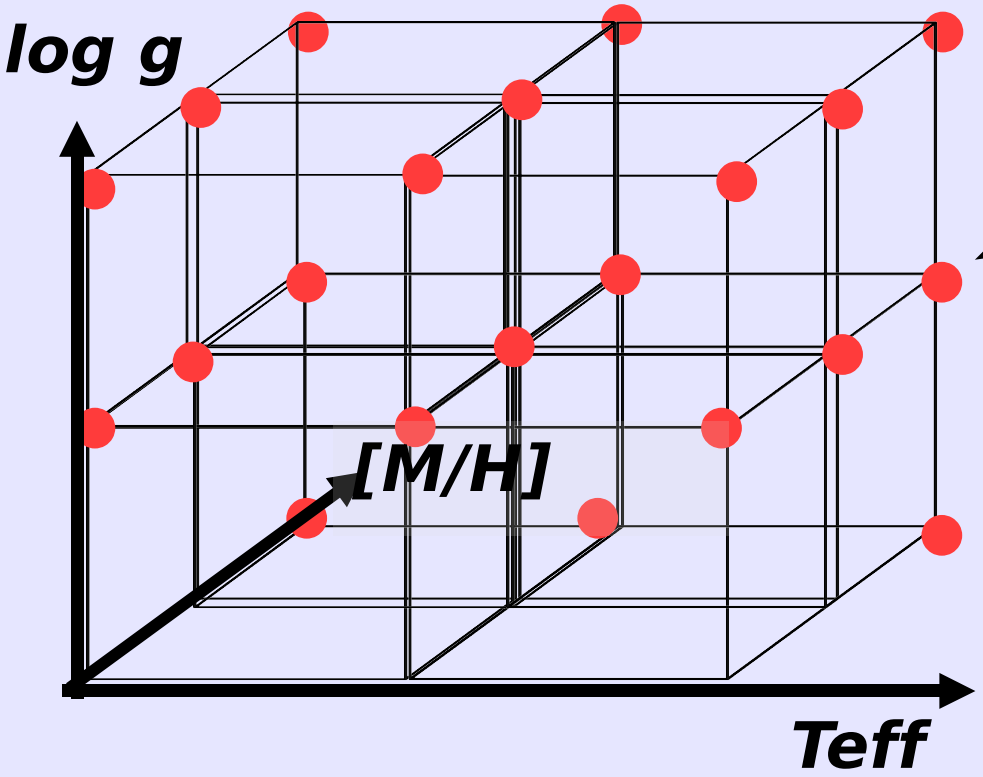
Impossible to work with analytical expressions ! =>

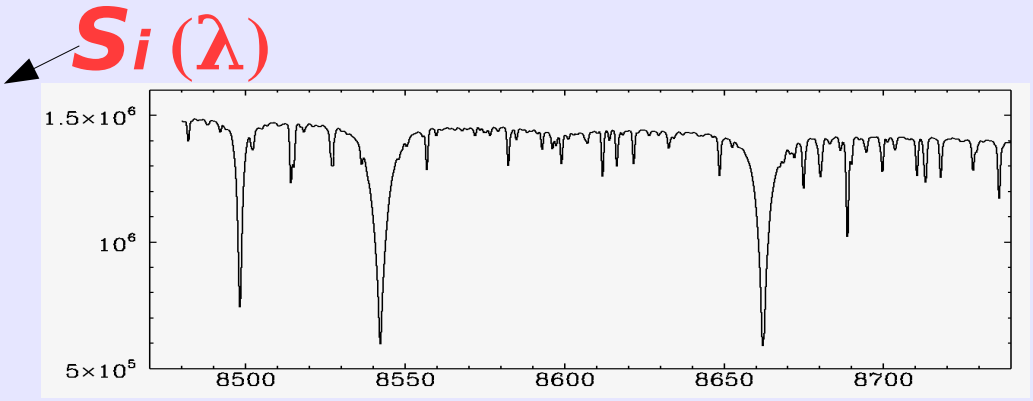
Grid of synthetic spectra

Automated stellar parametrization

2. Why is it difficult?

The parameter space (and the distance function) can suffer from :





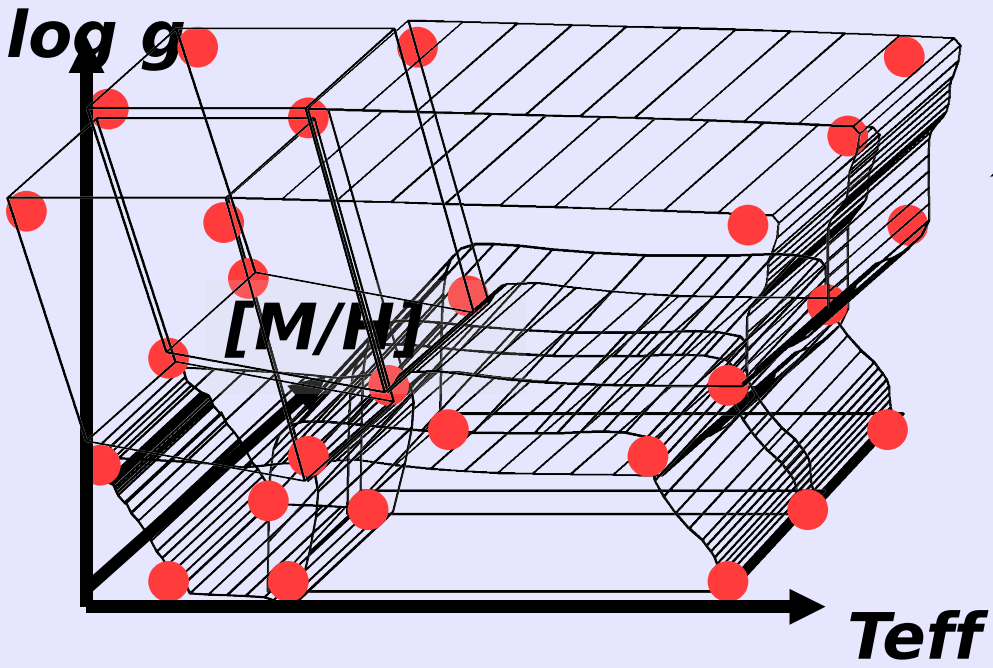
Automated stellar parametrization

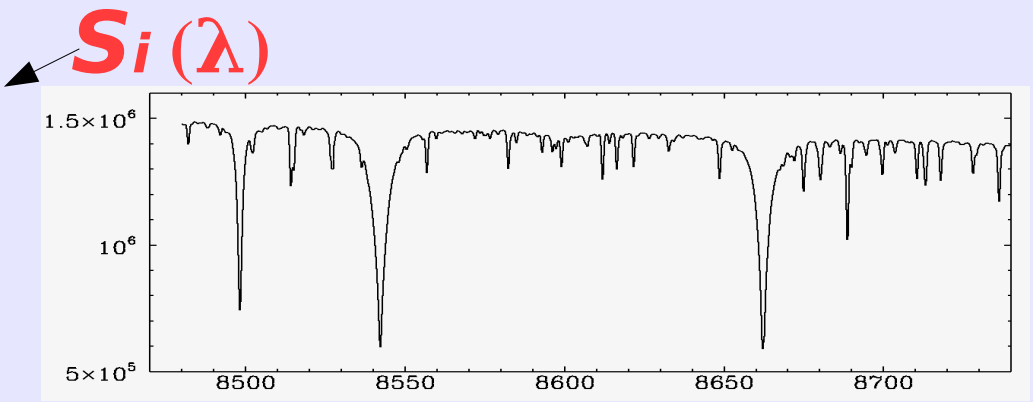
2. Why is it difficult?

The parameter space (and the distance function) can suffer from :

- **Non linearity : non-linear variations of the spectral flux with atmospheric parameters variations.**

  More important for large scales in the parameter space





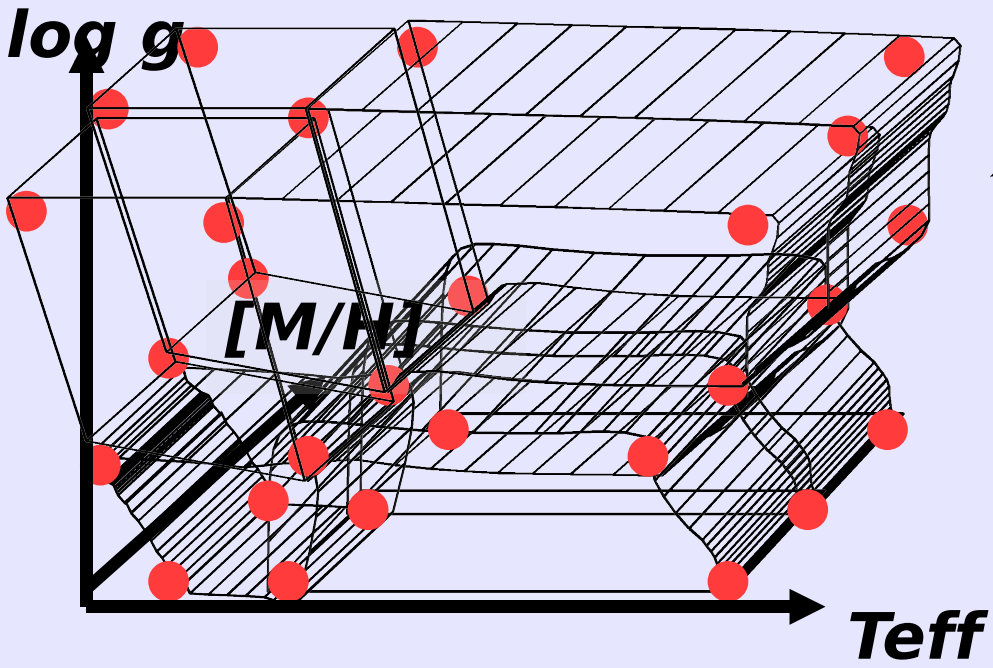
Automated stellar parametrization

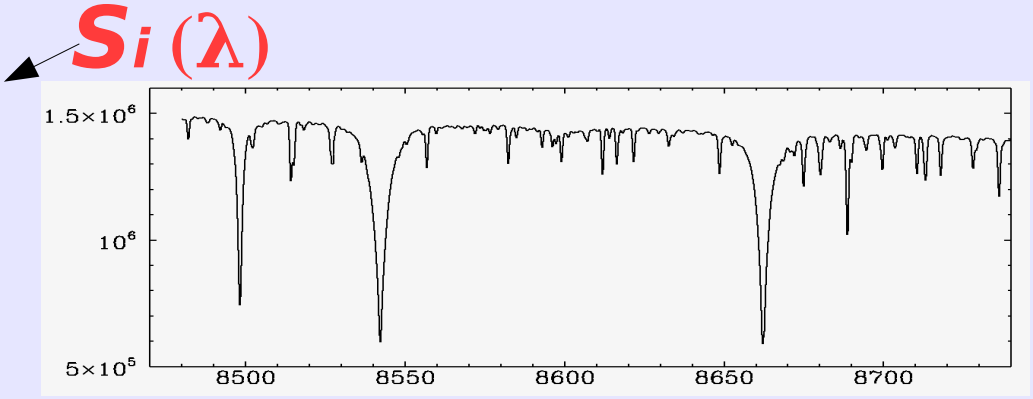
2. Why is it difficult?

The parameter space (and the distance function) can suffer from :

- **Non linearity**
- **Non convexity**

} **Degeneracy of the parameters :**





Automated stellar parametrization

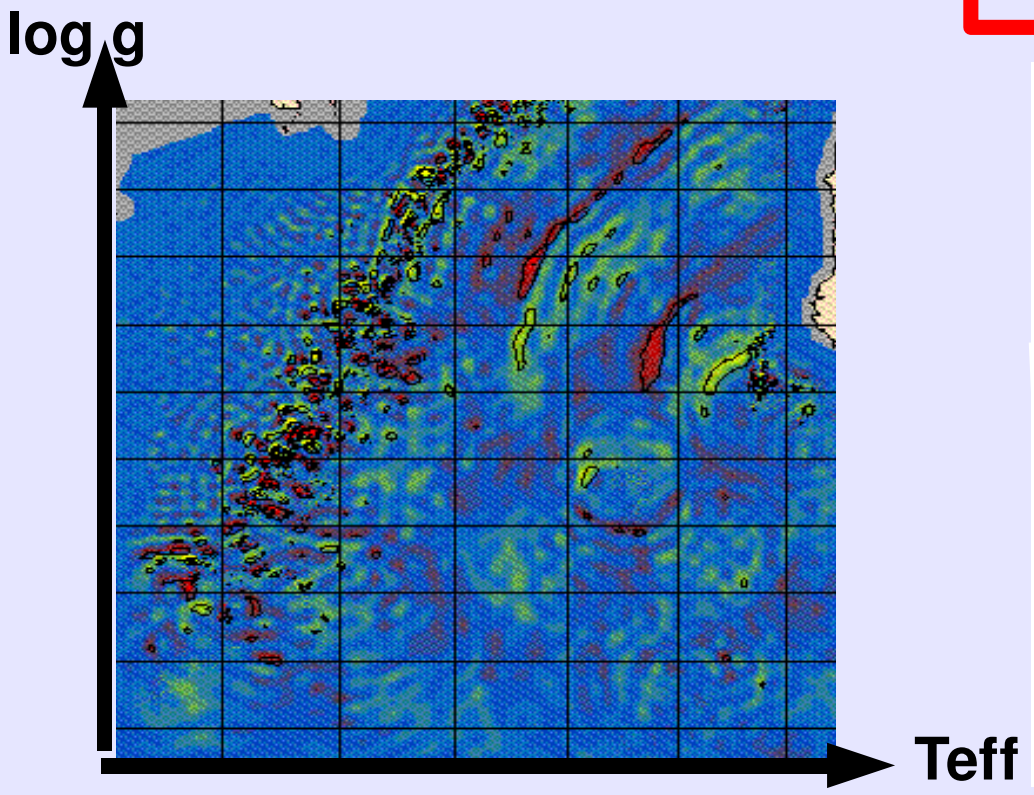
2. Why is it difficult?

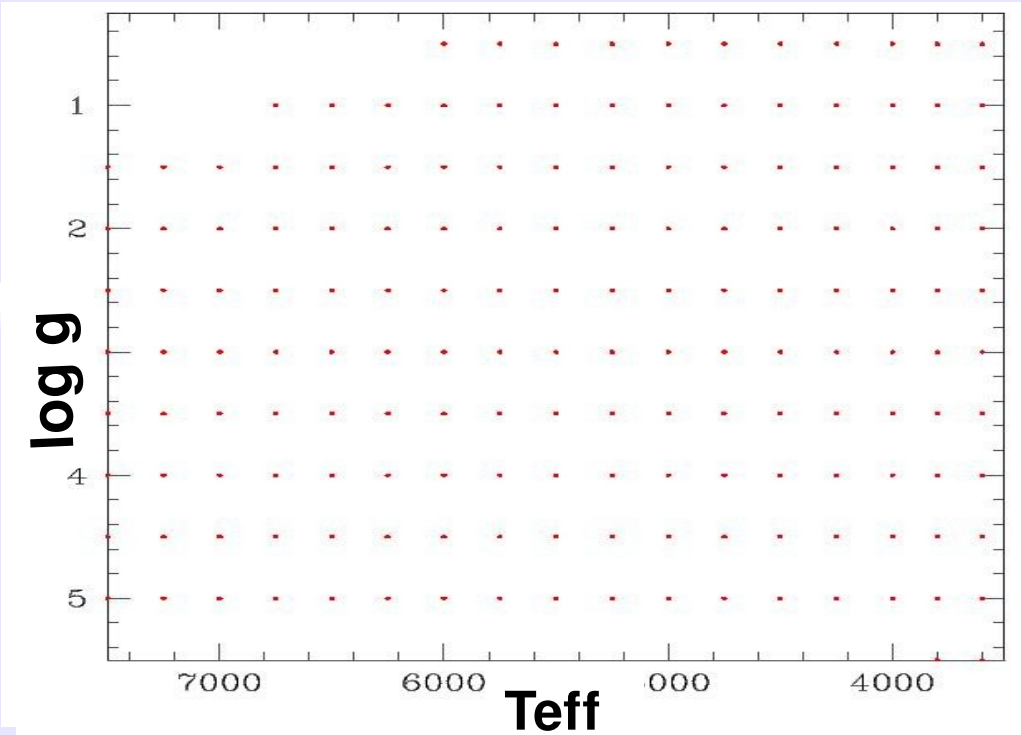
The parameter space (and the distance function) can suffer from :

- **Non linearity**
- **Non convexity**

Degeneracy of the parameters :

Secondary minima





Automated stellar parametrization

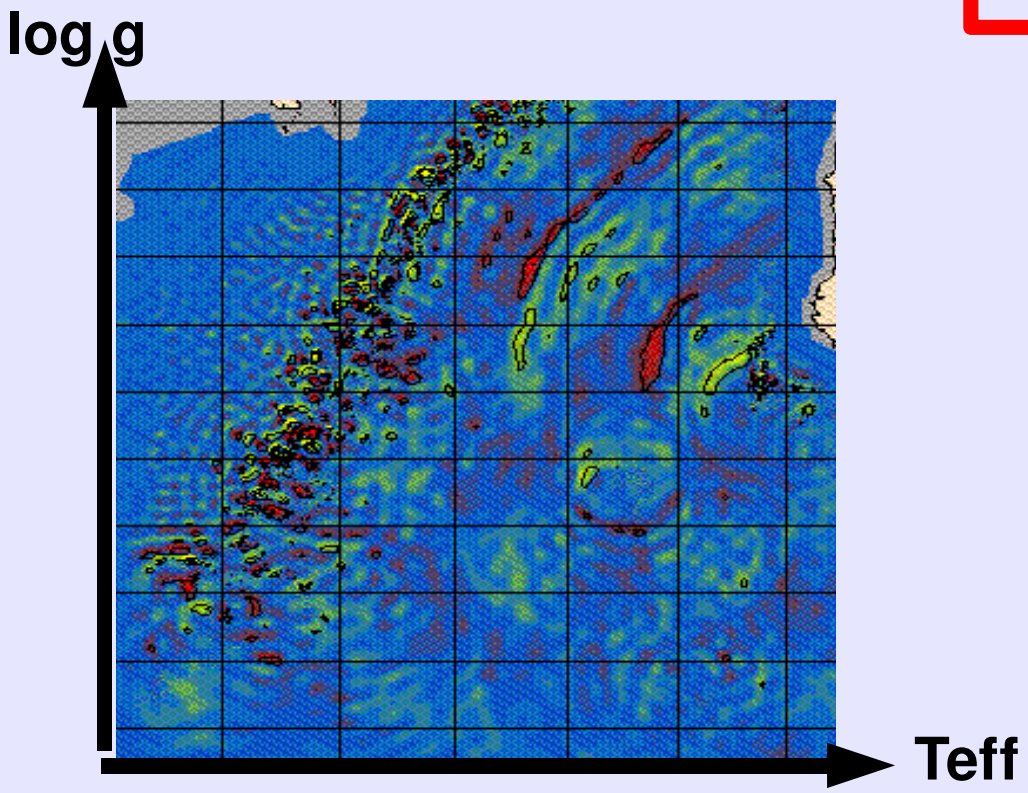
2. Why is it difficult?

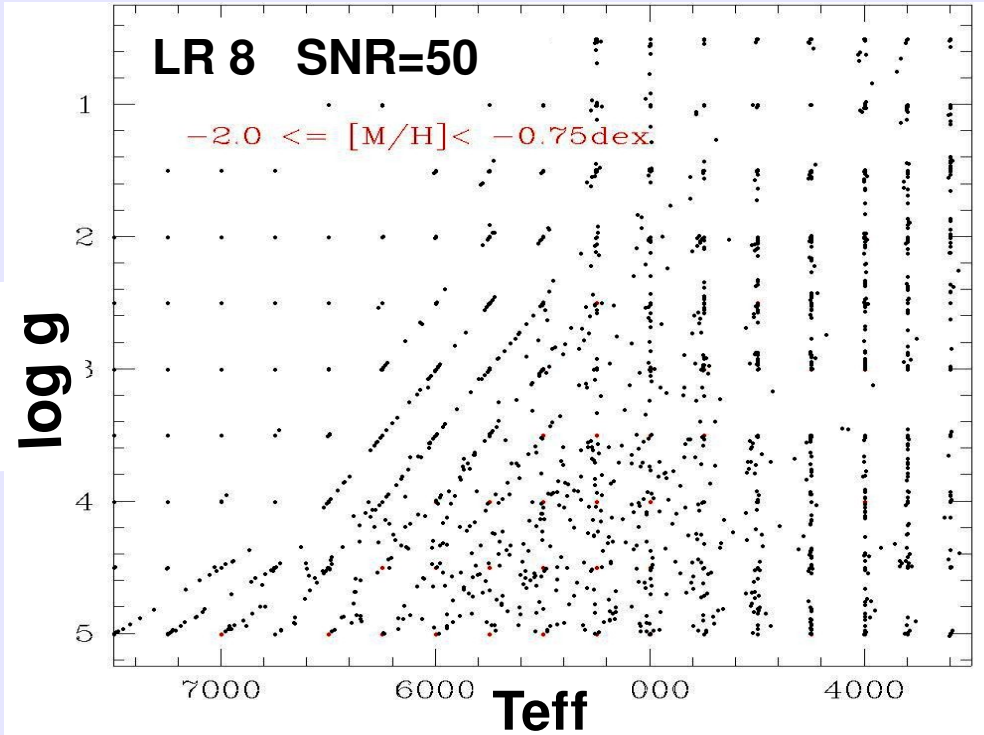
The parameter space (and the distance function) can suffer from :

- **Non linearity**
- **Non convexity**

Degeneracy of the parameters :

Secondary minima



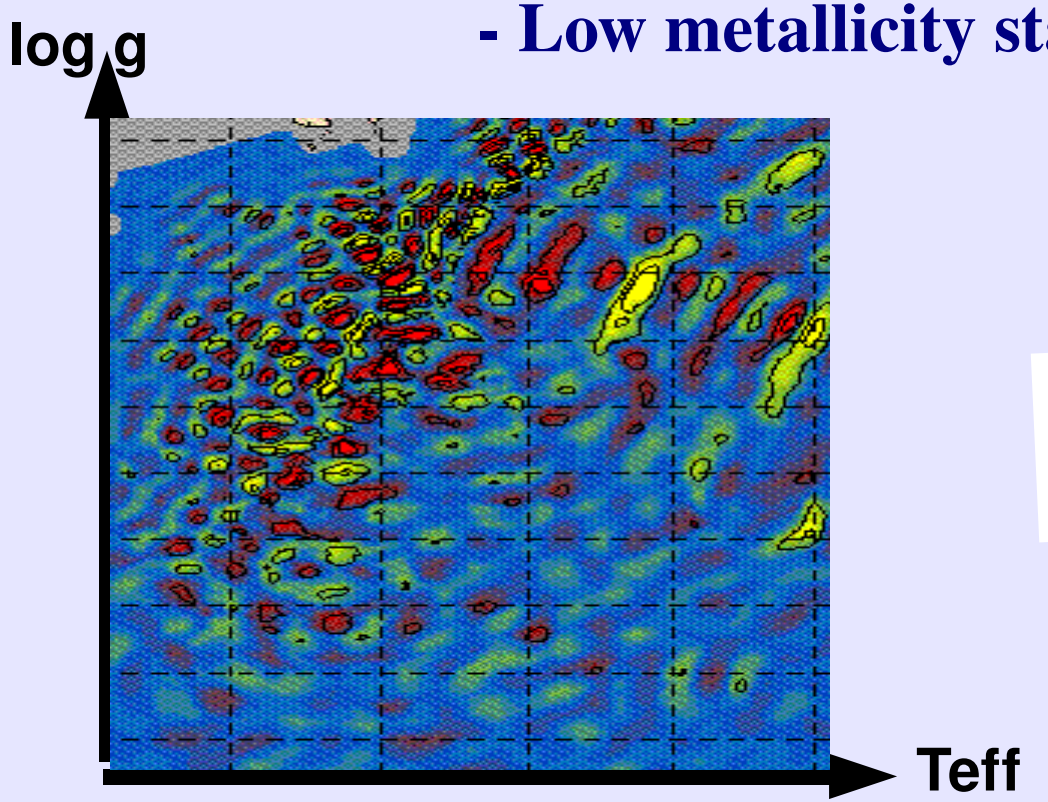


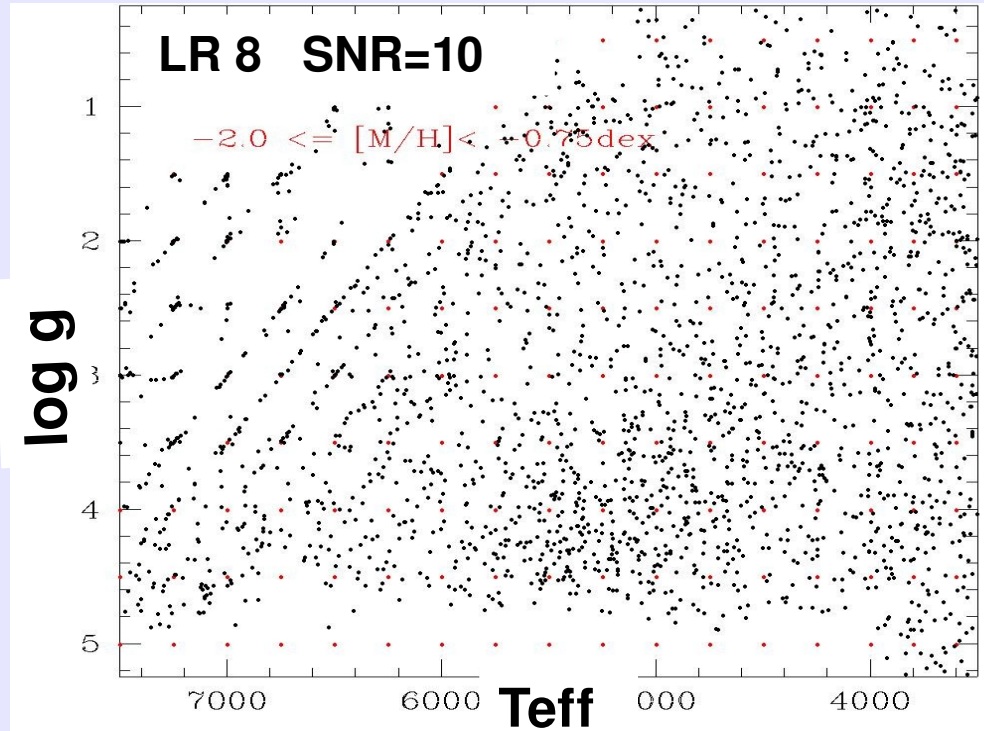
Automated stellar parametrization

2. Why is it difficult?

The situation gets worse when the spectra contain less information:

- **Low signal to noise ratio**
- **Low resolution**
- **Smaller spectral range**
- **Low metallicity stars, ...**





Automated stellar parametrization

3. Types of methods

Different mathematical approaches:

- **Optimization : parameters derived through a distance minimization**
- **Projection : spectra are projected into a set of vectors derived during a learning phase**
- **Classification : pattern recognition problem**

Automated stellar parametrization

3. Types of methods

1. Optimization methods

- Minimum distance algorithm : **all the distances are computed.**
  - **No problems with secondary minima**
  - **The precision is limited by the grid parameter sampling**
  - **Very time consuming**

- Nelder-Mead algorithm : **transformation of a simplex in the distance space by reflexion, contraction...**
  - **Succesfully applied to SDSS/SEGUE: Allende-Prieto et al. (2008)**
  - **Problems in a non-convex space : secondary minima traps**

Automated stellar parametrization

3. Types of methods

1. Optimization methods

- Gauss-Newton algorithm and its variants

  - ILIUM : **partial derivatives estimated by linear interpolations**
    Tested on simulated Gaia BP/RP data. Bailer-Jones (2010)

     **Problems in a non-convex space**

2. Projection methods

- Penalized chi square : **local linear model interpolations with regularization.**

   **Applied to RAVE (2nd release) : Zwitter et al. (2008)**

   **Problems in a non-convex space**

3. Types of methods

2. Projection methods

- MATISSE : **projection vectors are linear combinations of synthetic spectra. Two steps procedure to tackle non-linearity.**

Recio-Blanco et al. (2006)

- **Easy physical interpretability**
- **Very fast application**
- **Applied to FLAMES/GIRAFFE: Gazzano et al. (2010)**

Kordopatis et al. (2010, in prep.)

Applied to ESO archive data: AMBRE project (P.I. P. de Laverny)

Worley et al., in prep (her presentation)

Integrated into the Gaia DPAC pipeline at CNES for RVS data

- **Problems in a non-convex space**

3. Types of methods

2. Projection methods

- Principal Component Analysis : **the spectra grid is considered as a set of statistical trials of Npix variables (fluxes).**

  - MOPED : **Fisher matrix projection vectors associated to a data compression. Heavens et al. (2003)**

  - MAχ : **Simplified version with gaussian noise. Jofre et al. (2010)**


Very fast application. Tested on SEGUE and UVES data


Possible problems with non-linearity

3. Types of methods

3. Classification methods

- Artificial Neural Networks and SVM : **back-propagation and multi-layer algorithms or classification with hyperplanes. Bailer-Jones (2000)**

  - **Very fast application**
  - **Difficult interpretability: black-box effect**

- Oblique k-decision trees (DichoDif) : **space-partitioning structure in a k-dim space with projection vectors. Bijaoui et al. (2010) OCA-Nice**

  - **Very fast and good results for non-convex, non-linear spaces**

    Applied to FLAMES/GIRAFFE data (Kordopatis et al. 2010, in prep)

Automated stellar parametrization

4. The Gaia context







Algorithm in charge of determining the stellar APs from RVS spectra of individual stars (CaII IR triplet, R = ~11 500/~6 000):
CU8/Generalized Stellar Parametrizer-spectroscopy (Recio-Blanco et al.)

Context (I) :

1. **Normalized and radial velocity corrected spectra**

2. **Preliminar selection of stellar spectra (Discrete Source Classifier)**

3. **Synthetic, simulated and standard spectra from DPAC**

Automated stellar parametrization

4. The Gaia context







Algorithm in charge of determining the stellar APs from RVS spectra of individual stars (CaII IR triplet, R = ~11 500/~6 000):
CU8/Generalized Stellar Parametrizer-spectroscopy (Recio-Blanco et al.)

Context (II) :

4. **Four algorithms tested : 2 of them integrated**
5. **Integration at CNES for DPAC in 6-months cycles**
   Algorithms coded in Java with common tools implemented:
   Gaia data-model, dictionary tool, main database
6. **Possible inputs from GSP-phot (APs determined from BP/RP)**

Automated stellar parametrization

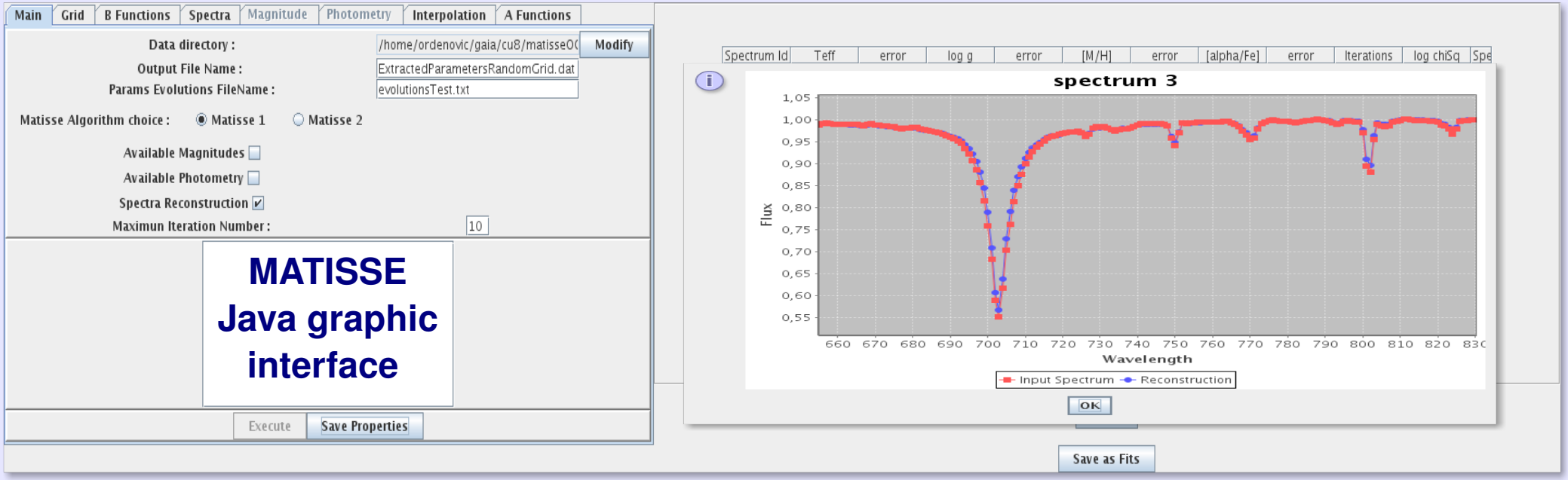
4. The Gaia context

Generalized Stellar Parametrizer-spectroscopy tested methods :

- **MATISSE**
- **Dicho-Dif**

} Observatoire de la Côte d'Azur (Nice) : A. Bijaoui, P. de Laverny, G. Kordopatis, Ch. Ordenovic, A. Recio-Blanco, C. Worley

Integrated at CNES: D-track delivery



- **Nelder-Mead**: C. Allende-Prieto (Instituto de Astrofísica Canarias)
- **Artificial Neural Networks**: D. Ordóñez (Univ. de A. Coruña)

Automated stellar parametrization

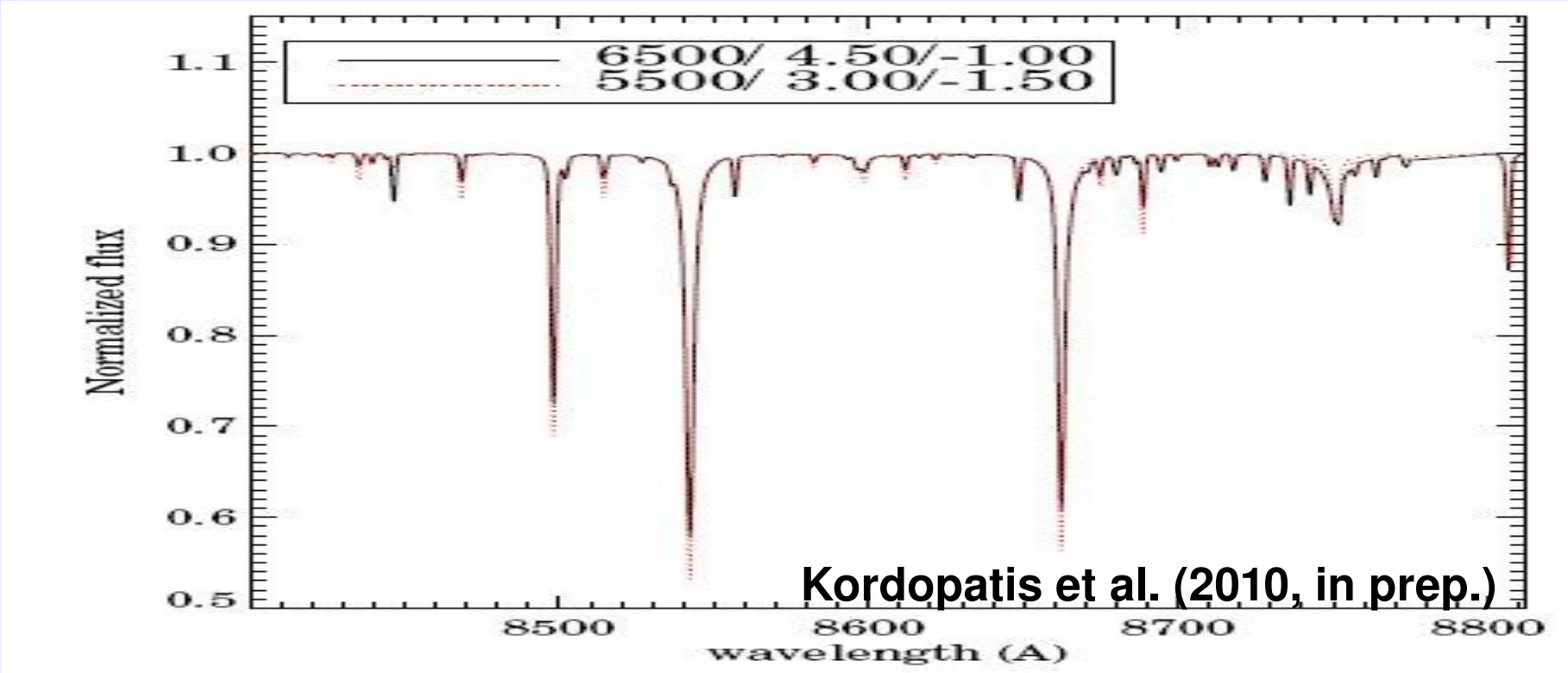
4. The Gaia context

Generalized Stellar Parametrizer-spectroscopy and the CaII range :

- **Internal errors :**

  Tests with HR and LR synthetic spectra (4 algorithms)

  Parameter degeneracies can be important (specially for LR)



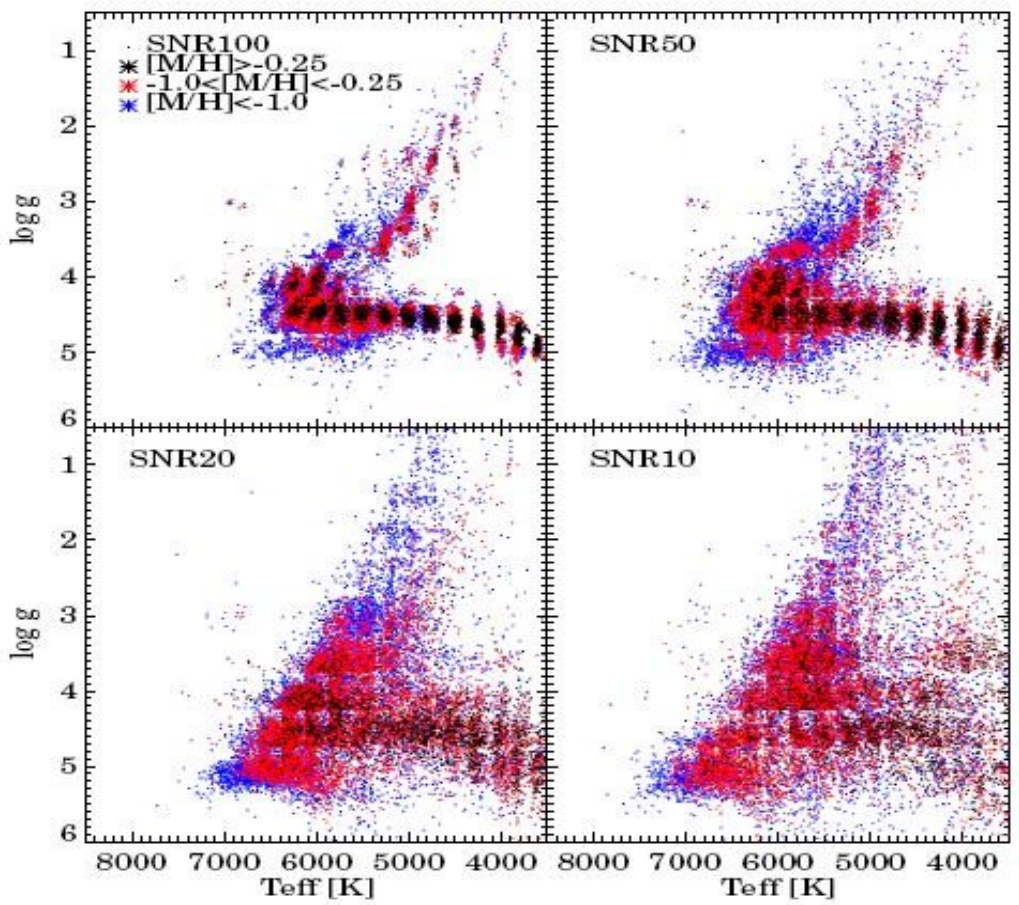
Automated stellar parametrization

4. The Gaia context

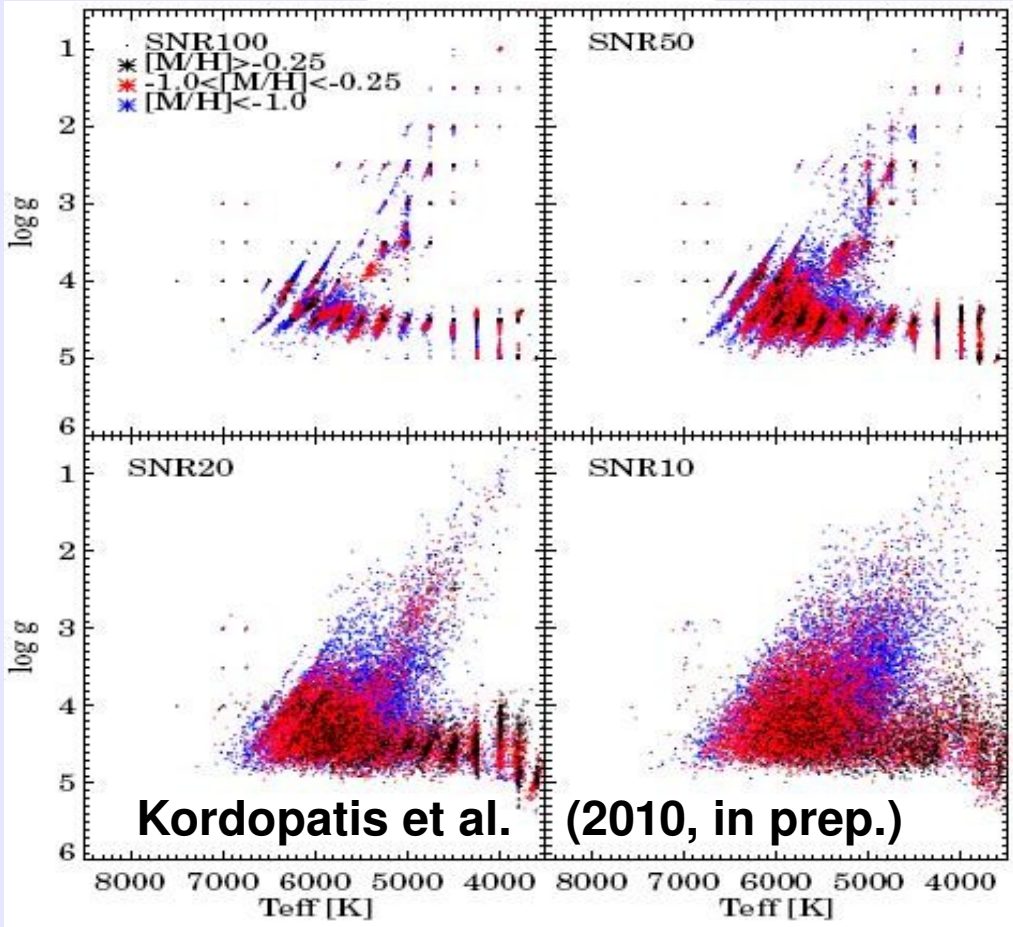
Generalized Stellar Parametrizer-spectroscopy and the CaII range :

Internal errors : Different methods depending on the SNR.

MATISSE



Dicho-Dif



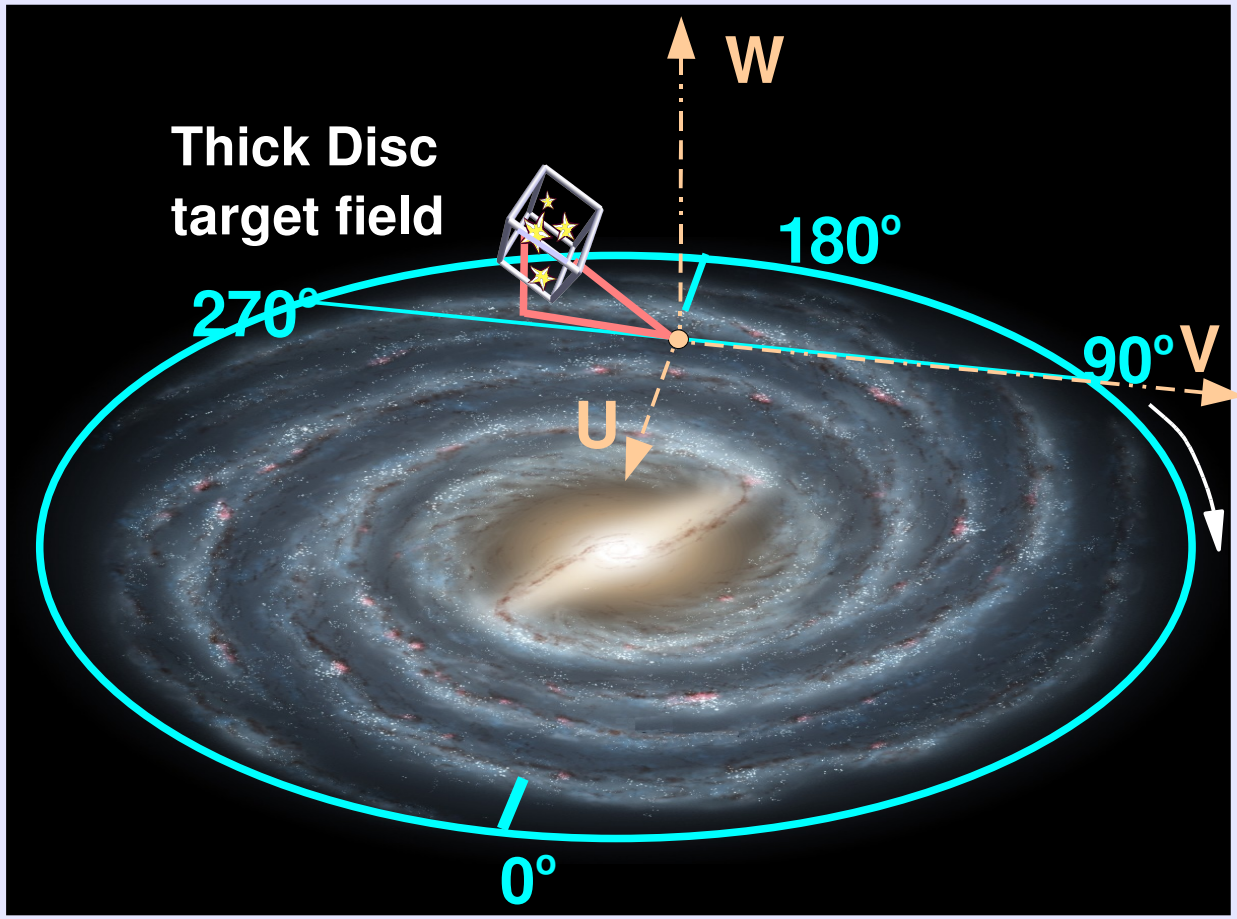
Automated stellar parametrization

4. The Gaia context

Generalized Stellar Parametrizer-spectroscopy and the CaII range :

Tests with real data : how to deal with real data problems ?

I. FLAMES LR8 data: G. Kordopatis PhD thesis



700 stars observed
dwarfs at ~ 2 kpc

Kordopatis et al. (2010, in prep.)

- **Teff, logg, [M/H]**
- **distances**
- **V, U, W**
- **angular momentum**
- **eccentricities**

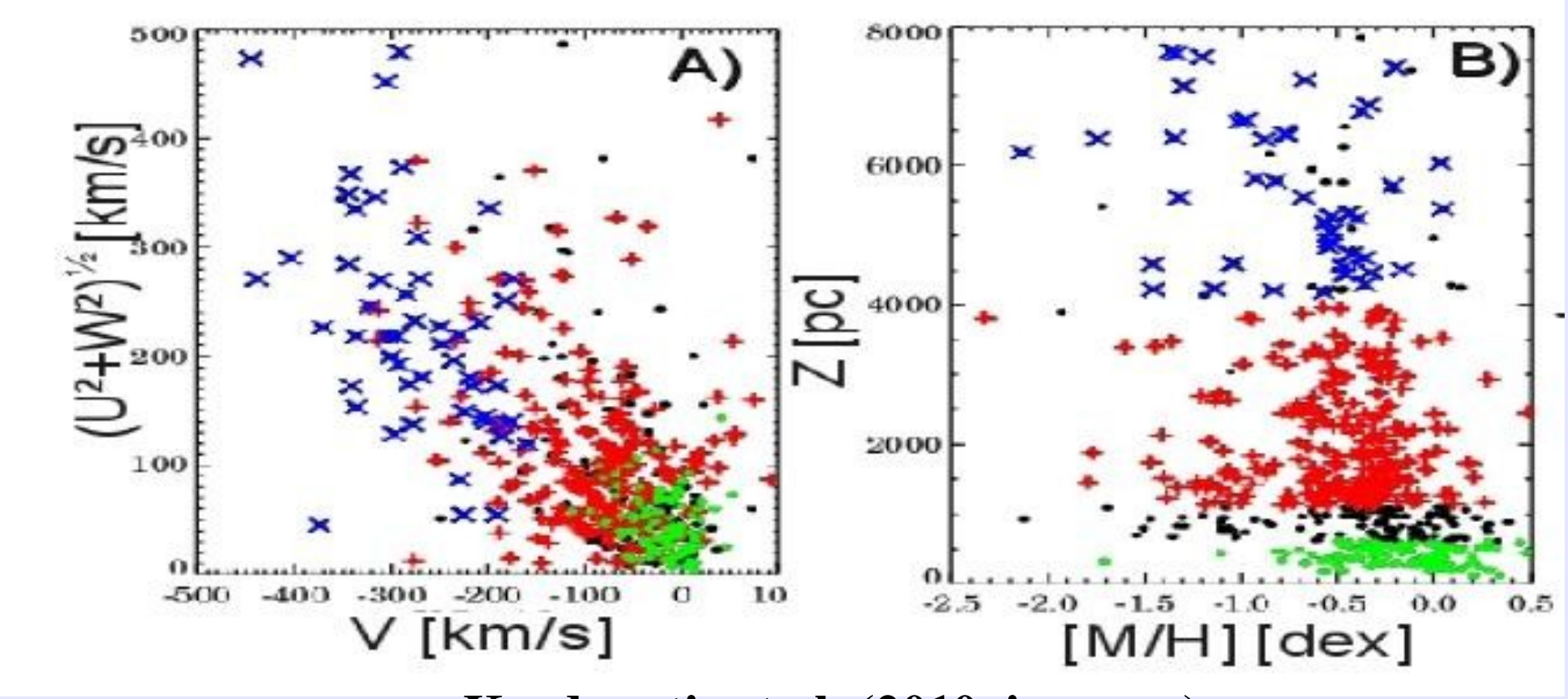
Automated stellar parametrization

4. The Gaia context

Generalized Stellar Parametrizer-spectroscopy and the CaII range :

Tests with real data : how to deal with real data problems ?

I. FLAMES LR8 data: G. Kordopatis PhD thesis



Kordopatis et al. (2010, in prep.)

Automated stellar parametrization

4. The Gaia context

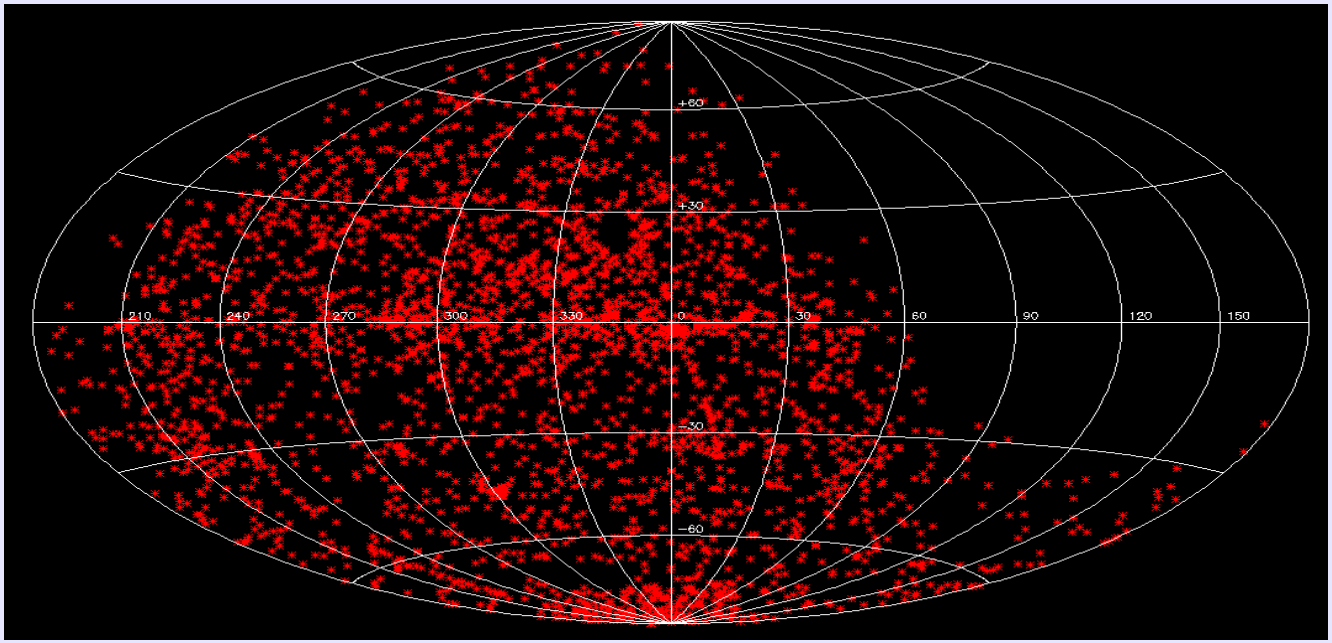
Generalized Stellar Parametrizer-spectroscopy and the CaII range :

Tests with real data : how to deal with real data problems ?

The AMBRE project (P.I. P. de Laverny, OCA)

Massive test with MATISSE : analysis of the ESO archive data:

FEROS, UVES, HARPS, FLAMES, see C. Worley's talk (15:30)



Galactic location of the FEROS stars : 22 000 spectra

Automated stellar parametrization

5. Conclusions

- **Several algorithms are ready to be applied to GCDS data**
  New algorithms are under-development or at different levels of testing / application ***(my list of examples was not exhaustive!)***

- **Each method has its application field, depending on the conditions of the parameter space for a given data set.**
  Several methods may be combined for an optimal result.

- **Individual chemical abundances :**
  - already existing methods adapted to an increasing number of parameters (ex. in Nice: DichoDif with PCA, MATISSE,...), EWs after atmosph. param. determination, SME, MyGIsFOS...

- **Gaia GSP-spec algorithm is an example of a communarity effort**