
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Vukadinoviac, D.; Smitha, H. N.; Korpi-Lagg, A.; Van Noort, M.; Castellanos Durán, J. S.; Solanki, S. K.

globin : A spectropolarimetric inversion code for the coupled inference of atomic line parameters

Published in:
Astronomy and Astrophysics

DOI:
[10.1051/0004-6361/202347752](https://doi.org/10.1051/0004-6361/202347752)

Published: 01/06/2024

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:
Vukadinoviac, D., Smitha, H. N., Korpi-Lagg, A., Van Noort, M., Castellanos Durán, J. S., & Solanki, S. K. (2024). globin : A spectropolarimetric inversion code for the coupled inference of atomic line parameters. *Astronomy and Astrophysics*, 686, 1-16. Article A262. <https://doi.org/10.1051/0004-6361/202347752>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

globin: A spectropolarimetric inversion code for the coupled inference of atomic line parameters

D. Vukadinović^{1,2}, H. N. Smitha¹, A. Korpi-Lagg^{1,3}, M. van Noort¹,
J. S. Castellanos Durán¹, and S. K. Solanki¹

¹ Max Planck Institute for Solar System Research, Justus-von-Liebig Weg 3, Göttingen, Germany
e-mail: vukadinovic@mps.mpg.de

² Georg-August-Universität Göttingen, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany

³ Department of Computer Science, Aalto University, PO Box 15400, 00076 Aalto, Finland

Received 18 August 2023 / Accepted 27 March 2024

ABSTRACT

Context. The reliability of physical parameters describing the solar atmosphere inferred from observed spectral line profiles depends on the accuracy of the involved atomic parameters. For many transitions, atomic data, such as the oscillator strength ($\log(gf)$) and the central wavelength of the line, are poorly constrained or even unknown.

Aims. We present and test a new inversion method that infers atomic line parameters and the height stratification of the atmospheric parameters from spatially resolved spectropolarimetric observations of the Sun. This method is implemented in the new inversion code `globin`.

Methods. The new method employs a global minimization algorithm enabling the coupling of inversion parameters common to all pixels, such as the atomic parameters of the observed spectral lines. At the same time, it permits the optimum atmospheric parameters to be retrieved individually for each spatial pixel. The uniqueness of this method lies in its ability to retrieve reliable atomic parameters even for heavily blended spectral lines. We tested the method by applying it to a set of 18 blended spectral lines between 4015 Å and 4017 Å, synthesized from a 3D magnetohydrodynamic simulation containing a sunspot and the quiet Sun region around it. The results were then compared with a previously used inversion method where atomic parameters were determined for every pixel independently (pixel-by-pixel method). For the same spectral region, we also inferred the atomic parameters from the synthesized spatially averaged disc-centre spectrum of the quiet-sun.

Results. The new method was able to retrieve the $\log(gf)$ values of all lines to an accuracy of 0.004 dex, while the pixel-by-pixel method retrieved the same parameter to an accuracy of only 0.025 dex. The largest differences between the two methods are evident for the heavily blended lines, with the former method performing better than the latter. In addition, the new method is also able to infer reliable atmospheric parameters in all the inverted pixels by successfully disentangling the degeneracies between the atomic and atmospheric parameters.

Conclusions. The new method is well suited for the reliable determination of both atomic and atmospheric parameters and works well on all spectral lines, including those that are weak and/or severely blended. This is of high relevance, especially for the analysis of observations of spectral regions with a very high density of spectral lines. An example includes the future near-ultraviolet spectropolarimetric observations of the SUNRISE III stratospheric balloon mission.

Key words. atomic data – methods: data analysis – methods: numerical – Sun: atmosphere

1. Introduction

The physical conditions in the solar atmosphere are encoded in the spectral lines, arising from absorption and emission processes in atoms, ions, and molecules. By modelling the observed spectral line profiles and their polarization states, we can retrieve a height stratification of the atmospheric parameters, such as temperature, magnetic field, and line-of-sight velocity.

To model the observed line profiles, the spectral lines are synthesized first with an initial guess model for the atmosphere and then compared with the observations. The differences between the observed and the synthetic spectra are iteratively reduced by adjusting the input atmospheric model used for the synthesis. The model that reproduces the observed spectrum is deemed to be a good representation of the atmospheric conditions of the region from which the observed spectrum

was emitted. This is the basic principle behind spectropolarimetric inversion techniques (e.g. Ruiz Cobo & del Toro Iniesta 1992; Frutiger et al. 2000; del Toro Iniesta & Ruiz Cobo 2016) that are routinely used to obtain an atmospheric model from the observed Stokes profiles. The retrieval of atmospheric parameters using inversions depends on the following: the height range over which a line is sensitive to the atmospheric parameters; the approximations used for the radiative transfer computations (e.g. de la Cruz Rodríguez & van Noort 2017; Smitha et al. 2020), how well one can correct for the contamination of Stokes profiles due to the cross talk between them; stray light or smearing by the telescope point spread function (PSF; e.g. Orozco Suárez et al. 2007; van Noort 2012); and, most importantly, the accuracy of the atomic line parameters used for spectral synthesis.

Together with the local atmospheric conditions, atomic parameters are an important factor in determining the spectral

line shape. The oscillator strength f is one of the key parameters that characterizes the spectral line strength. The parameter f is a quantum correction to a classically derived transition probability between two energy levels in an atom, ion, or molecule (Mihalas 1978). This parameter is usually given in combination with the statistical weight, g , of the lower transition level, as $\log(gf)$ in atomic line databases. The unit measure for $\log(gf)$ is dex (decimal exponent).

Poorly determined atomic parameters remain a problem in spectroscopy, with the most important atomic parameters for studies carried out in local thermodynamic equilibrium (LTE) being the excitation potential of the lower energy level of the transition, the central wavelength of the spectral line, $\log(gf)$. The $\log(gf)$ value for a single spectral line can vary greatly between different atomic line databases such as that of the National Institute of Standards and Technology (NIST; Kramida et al. 2022), Kurucz (Kurucz et al. 1995) or the Vienna Atomic Line Database (VALD; Piskunov et al. 1995). The origin of these differences mainly lies in the methods used for their determination: laboratory measurements using a spectroscopic furnace (e.g. Blackwell & Collins 1972) or laser-induced fluorescence (e.g. Den Hartog et al. 2005), theoretical computations based on an atomic model (e.g. Pradhan & Saraph 1977), or using observed solar spectra (e.g. Gurtovenko & Kostik 1981, 1982; Thévenin 1989, 1990; Borrero et al. 2003; Bigot & Thévenin 2006; Shchukina & Vasil'eva 2013; Trelles Arjona et al. 2021).

Laboratory measurements provide precise $\log(gf)$ values only for isolated and strong lines. Theoretical computations exclusively depend on the distribution of energy levels in the atomic model and, therefore, have no restrictions on line strength. However, already a small error in the calculation of the energy levels in the atomic model leads to a change in parameters such as the central wavelength of the line, collisional rates, and consequently to a large uncertainty of the $\log(gf)$ values (Pradhan & Saraph 1977). The advantages and pitfalls of every method are summarised in detail in Borrero et al. (2003) and Shchukina & Vasil'eva (2013). The present work focusses on the inference of atomic parameters based on spatially resolved solar spectra.

An observed solar spectrum permits the inference of atomic parameters of lines whose profiles are unattainable in laboratory setups due to the difficulty in reproducing the physical conditions prevailing in the solar atmosphere. Thus, observed solar spectra provide a unique way of inferring the properties of spectral lines. There are two inversion-based methods commonly used to determine the $\log(gf)$ from the solar spectrum. One is the spatially averaged method, and the other is the pixel-by-pixel method. In the first method, the $\log(gf)$ inference involves modelling the observed spatially averaged solar spectrum with a synthetic spectrum computed using a mean solar atmospheric model by the iterative adjustment of $\log(gf)$ values (Gurtovenko & Kostik 1981, 1982; Thévenin 1989, 1990; Borrero et al. 2003; Shchukina & Vasil'eva 2013, and references therein). The reliability of the inferred $\log(gf)$ values using this method depends on the quality of the used atmospheric model, adopted abundances of atomic elements, and proper treatment of line broadening to model the observed line profiles accurately. For example, Bigot & Thévenin (2006) used the temporally and spatially averaged spectra from a 3D hydrodynamic model of the solar atmosphere to derive $\log(gf)$ from the observed spatially averaged solar spectrum, thus removing the problem of choosing the correct atmospheric model and macro- and micro-turbulent broadening.

Borrero et al. (2003) showed that uncertainties in the central wavelength of a line ($\Delta\lambda$) must be accounted for in the spatially averaged method for a reliable retrieval of $\log(gf)$. They estimated the uncertainty in the central wavelength from the spatially averaged method is 5 mÅ at 1 μm , which is comparable to the uncertainty achieved in laboratory measurements of Fe I lines (Nave et al. 1994). Precise central wavelengths of spectral lines are of significant importance in revising the energy levels of atoms and ions (Borrero et al. 2003), identifying new energy levels in atoms and ions (Peterson & Kurucz 2022), and for accurate identification of spectral lines in line-rich spectra, especially of iron-group elements (Nave et al. 2017).

The pixel-by-pixel method is applied to spatially resolved spectra. With this method, the atomic parameters are inferred independently for each pixel within the observed field-of-view (for further description of this method, see Appendix A). Trelles Arjona et al. (2021) use the pixel-by-pixel method to simultaneously infer the atmospheric parameters along with $\log(gf)$ values for 15 spectral lines around 1.56 μm . The final $\log(gf)$ value is the average of all $\log(gf)$ values retrieved from pixels with the best match of the synthetic spectrum to the observed spectrum. The strong cross talk between atmospheric and atomic parameters can result in a very diverse pool of $\log(gf)$ values, thereby introducing uncertainties in both the inferred atomic as well as atmospheric quantities.

In the present work, we develop a new inversion method for a reliable determination of both atmospheric as well as atomic parameters for any number of spectral lines simultaneously. During inversions, the atomic parameters for a given spectral line are not allowed to vary at each pixel independently, like in the pixel-by-pixel method, but are allowed to vary only simultaneously in all the pixels. At the same time, this new inversion method allows for the retrieval of the atmospheric parameters in each pixel independently. This new method is part of the new inversion code named *globin*, and we refer to this new method in the further text as the coupled method.

The comparison of the spatially averaged observed spectrum and the average synthetic spectrum from a 3D atmosphere shows differences in spectral line strengths in NUV, resulting mainly from poor knowledge of atomic parameters (Riethmüller & Solanki 2019). To retrieve reliable atmospheric parameters from NUV spectra, we first have to determine the atomic parameters of the spectral lines accurately. The high spectral line density (especially in the NUV) causes the absorption profiles to overlap. This so-called blending of spectral lines is an additional complication for the reliable retrieval of the often poorly known atomic parameters of these lines and the atmospheric parameters. We expect that the coupled method will be able to provide a reliable inference of inversion parameters despite the heavy blending between spectral lines.

This work is motivated by the future high spatial and spectral resolution observations in NUV from the instrument SUNRISE UV Spectropolarimeter and Imager (henceforth SUSI, Feller et al. 2020) onboard the next flight of the balloon-borne SUNRISE observatory which has successfully flown twice before (Solanki et al. 2010, 2017) and is expected to fly again in June 2024. SUSI will continuously observe the Sun in the near-ultraviolet (NUV) spectral range from 309 nm to 417 nm. Radiation at these wavelengths is strongly absorbed by the Earth's atmosphere, which makes it hard to observe them from the ground.

The application of the coupled method is not restricted to the spectral region observed by SUSI. This method is general enough to be applied to spectral lines in other wavelength

regions as well. Additionally, this method is not limited to the inference of only $\log(gf)$ and $\Delta\lambda$. It can easily be extended to infer, for example, elemental abundances. However, such an extension is outside the scope of this work and will be addressed in future studies.

Efforts have been made to compile a list of spectral lines that have reliable atomic parameters (e.g. [Bigot & Thévenin 2006](#); [Heiter et al. 2021](#)) to determine the chemical composition of stellar atmospheres observed by the *Gaia* mission ([Gaia Collaboration 2016](#)). Similar to the coupled method that uses spectra from different pixels to constrain the atomic parameters, spectra from different stars have been used instead to infer the atomic parameters of lines predominantly in the visible and infrared wavelengths (e.g. [Boeche & Grebel 2016](#); [Martins et al. 2014](#); [Laverick 2019](#)). We expect that the coupled method would aid in providing reliable atomic parameters of spectral lines that would also be useful to the wider stellar physics community.

Section 2 introduces the coupled method, and in Sect. 3, we demonstrate the potential of this method for inferring $\log(gf)$ and to correct for any errors in the available central wavelength of the spectral lines. The results from the coupled method are compared to those from the pixel-by-pixel method. For this comparison, we use synthetic multi-line spectra computed from a realistic 3D magnetohydrodynamic (MHD) simulation of the solar atmosphere made with the MPS/University of Chicago Radiative MHD code (MURaM; [Rempel 2012](#)). In Sect. 4, we apply the spatially averaged method to a spatially averaged synthetic spectrum from a quiet-sun MURaM atmospheric model. Here, we test the applicability of the spatially averaged method for the retrieval of atomic parameters from blended spectral lines. Conclusions are given in Sect. 5.

2. The coupled method

Spectropolarimetric inversion is a degenerate problem where different sets of atmospheric parameters can result in very similar spectra (for a review see [del Toro Iniesta & Ruiz Cobo 2016](#)). This implies that a certain degree of cross talk exists between the atmospheric parameters, even when the atomic parameters are well known. When the atomic parameters are also inverted, this cross talk becomes even larger. It is especially pronounced between those quantities that affect a given spectral line in a similar way, for example, between temperature and $\log(gf)$, and between line-of-sight velocity and uncertainties in the central wavelength of the line.

To circumvent this problem, we have developed the coupled method, where the inference of atmospheric parameters is done individually at every pixel while fitting for only one set of atomic parameters for all pixels in a coupled way. The coupled method uses the Levenberg-Marquardt (LM) optimization algorithm for the inference of the atmospheric and atomic parameters. In the LM optimization scheme, we have to solve the equation:

$$\mathcal{H}\Delta\mathbf{p}_j = \mathcal{J}^T \cdot \Delta, \quad (1)$$

$$\text{where } \Delta_i = \sqrt{\frac{2}{N-n}} \cdot \frac{w_i}{\sigma_i} \cdot (\mathbf{O}_i - \mathbf{S}_i(\mathbf{p}_j))^1.$$

The solution of Eq. (1) gives the correction $\Delta\mathbf{p}_j$ to the current set of parameters \mathbf{p}_j . The correction to the atomic parameters is determined from differences between observed \mathbf{O} and computed spectra $\mathbf{S}(\mathbf{p}_j)$ in all pixels. In contrast, the corrections to the atmospheric parameters of individual pixels are determined from the differences between spectra in that pixel alone. This

kind of parameters' correction is achieved by a separation of the atmospheric and atomic parameters in the Jacobian matrix of the system \mathcal{J} as:

$$\mathcal{J} = \left(\text{diag}(\mathcal{J}_1^{\text{atm}}, \dots, \mathcal{J}_N^{\text{atm}}) \mid \mathcal{J}^{\text{atom}} \right) \quad (2)$$

where $\mathcal{J}_i^{\text{atm}}$ is the Jacobian matrix of the atmospheric parameters in the pixel i ($i = \overline{1, N}$) from the observed field of view, and the $\mathcal{J}^{\text{atom}}$ is the Jacobian matrix of the atomic parameters (spanning all rows in \mathcal{J}). The columns of these Jacobian matrices correspond to the response functions of the atmospheric and atomic parameters that describe how the spectrum is altered by changing the inversion parameters. The Hessian matrix \mathcal{H} is computed from the Jacobian matrix as $\mathcal{H} = \mathcal{J}^T \mathcal{J} + \lambda_M \cdot \text{diag}(\mathcal{J}^T \mathcal{J})$, where λ_M is the Marquardt parameter that controls the sampling of the parameters space.

By coupling the atomic parameters for all pixels, the total number of (free) inversion parameters is reduced compared to the pixel-by-pixel method, in which they vary independently for each pixel. The coupling of atomic parameters constrains the sampling of the parameter space by the optimization method and aids in obtaining the correct set of inversion parameters. A more detailed description of the coupled method is given in Appendix B.

The idea of coupling the spectra in different pixels originates from the 2D inversion scheme that was proposed in [van Noort \(2012\)](#). There, the author uses the telescope PSF to couple the spectra of nearby pixels in order to retrieve inversion parameters. This improves the spatial resolution of the inferred atmosphere while also reducing the impact of noise on the inferred parameter values (see e.g. [van Noort et al. 2013](#); [Tiwari et al. 2013, 2015](#); [Castellanos Durán et al. 2023](#)). Additionally, [de la Cruz Rodríguez \(2019\)](#) introduces a spatial regularization method limiting the spatial variation of desired inversion parameters. Both of these techniques are concerned mainly with retrieving atmospheric parameters, whereas in the present work, we focus on atomic line parameters. We chose to follow the approach of [van Noort \(2012\)](#) and extend the 2D-coupling scheme to fit any type of global, field-of-view independent parameters, in this paper, the atomic parameters' $\log(gf)$ and $\Delta\lambda$.

This new inversion method is implemented in the `globin` inversion code written in Python that uses the RH code ([Uitenbroek 2001](#)) for the spectral synthesis². RH solves the polarized radiative transfer equation for non-LTE (NLTE) line formation using the multilevel accelerated lambda iteration scheme ([Rybicki & Hummer 1991](#)). We opted for the RH code because of its versatility in treating NLTE line formation, partial frequency redistribution and continuum opacity fudge correction ([Bruls et al. 1992](#)). All these features will be necessary for future analyses, such as NUV spectra from the SUSI instrument. A detailed description of the implementation and functionalities of `globin` are given in Appendix C.

3. Comparison between the pixel-by-pixel method and the coupled method

To test the capabilities of the coupled method in comparison to the pixel-by-pixel method, we analysed synthetic Stokes profiles

² The code was adapted to be imported in a Python environment as a pre-compiled module. All the functionalities of the original code are preserved.

¹ For a comprehensive derivation of the Eq. (1), see Appendix A.

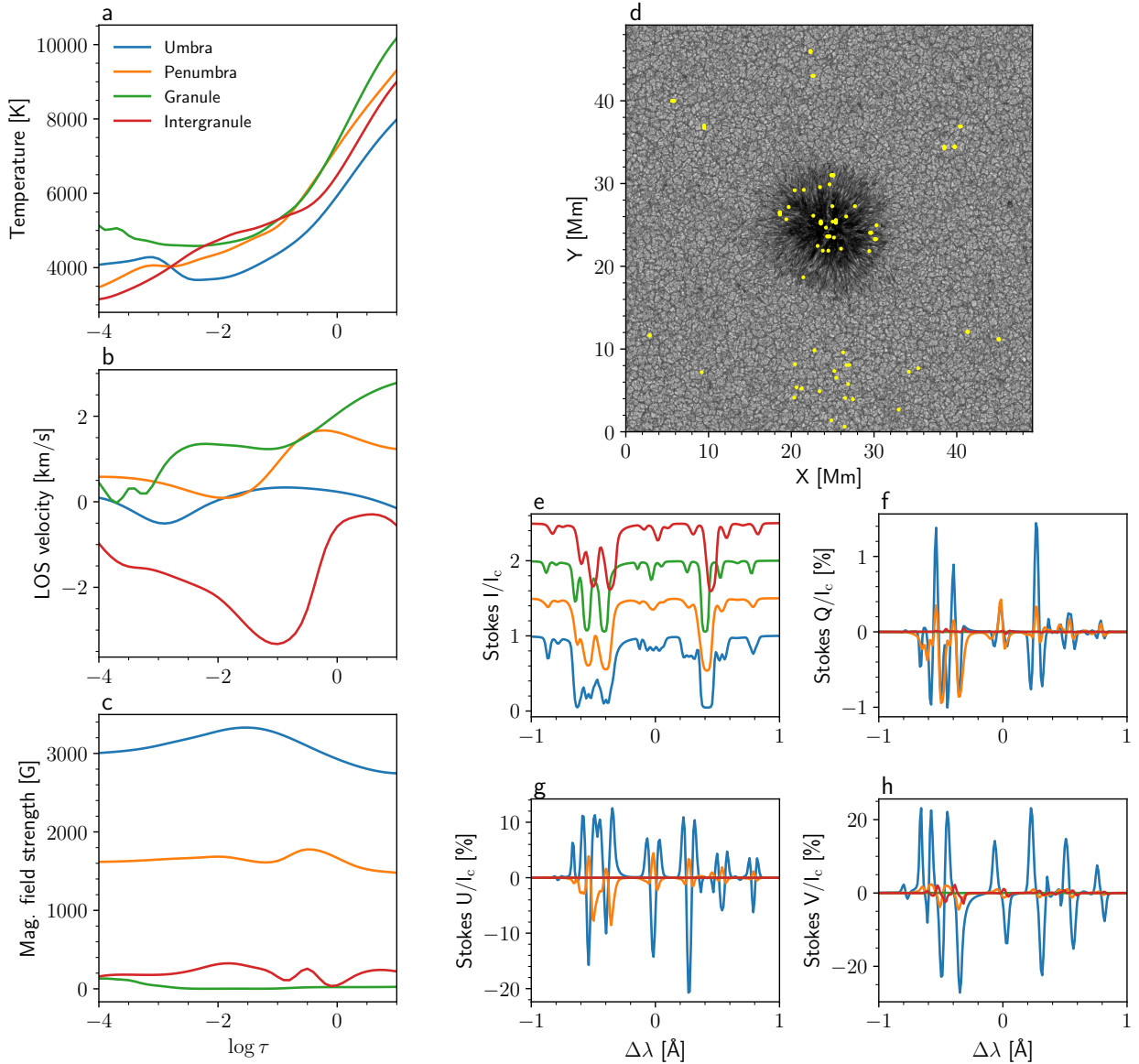


Fig. 1. Sample of Stokes spectra synthesized from the extracted atmospheric models of different solar features. Panels a-c: stratification of the temperature (a), LOS velocity (b), and magnetic field strength (c) from one pixel for each atmospheric feature taken from the sample of extracted atmospheres. Panel d: the continuum intensity from the MHD cube at 4015 Å with extracted atmospheres annotated with yellow dots. Panels e-h: Stokes parameters in the 4015–4017 Å spectral range from the atmospheres displayed in panels a-c. Wavelengths are given with respect to 4016 Å. For easier comparison, the Stokes profiles are normalized to the local continuum intensity I_c and the Stokes I/I_c profiles from different features are shifted vertically.

sampling different features such as an umbra, penumbra, granule and intergranular lane, obtained from a 3D MHD snapshot of a sunspot (Rempel 2012, see the continuum image in Fig. 1d) simulated using the MURaM code (Vögler et al. 2005). This selection guarantees a diverse sample of temperature, line-of-sight (LOS) velocity and magnetic field stratifications (Fig. 1a-c) resulting in very different Stokes profiles (Fig. 1e-h). This diversity leads to line profiles, especially of weaker lines, that are prominent in some features but not visible in others. In particular, the response of lines to a change in the atomic parameters is different for the different atmospheres. This result is beneficial for reducing the cross talk between atomic and atmospheric parameters in the coupled method.

We tested the applicability of the pixel-by-pixel and the coupled methods to determine the atomic parameters $\log(gf)$ and to

correct for possible errors in the central wavelength ($\Delta\lambda$) of spectral lines in the 4015–4017 Å range, containing several blended lines. This spectral region was chosen because it is covered by the Hamburg atlas spectrum (Neckel & Labs 1984), it has many blended spectral lines, and it lies within the wavelength range covered by the SUSI instrument. The same spectral range was considered when we tested the spatially averaged method (see section 4).

Spectral line information for the considered spectral region were taken from the Kurucz line list (Kurucz et al. 1995). There are many lines in the Kurucz line list that are fairly weak and are expected to have an insignificant effect on the spectrum. Including all of them will only extend the computing time. We considered only those spectral lines with a line core depth of at least 1% of the local continuum intensity. We found 18 such

Table 1. Atomic line parameters from the 4015–4017 Å range.

No	Element	λ_0 [Å]	E_i [eV]	g_{eff}	A	b [%]	D_{core}	$\log(gf)_0$	$\log(gf)_{\text{mode2}}$	$\log(gf)_{\text{mode3}}$	$\log(gf)_{\text{spat}}$
1	Fe I	4015.139	4.1631	0.666	7.44 ^(b)	3.9	0.149	-2.172	-2.087	-2.170	-2.166
2	Co I	4015.219	2.7849	0.484	4.92 ^(a)	32.9	0.033	-1.806	-1.646	-1.795	-1.610
3	Ti I	4015.373	2.0850	0.515	4.99 ^(a)	18.5	0.428	-0.084	-0.217	-0.069	-0.105
4	VI	4015.398	2.0618	0.995	4.00 ^(a)	98.3	0.007	-0.857	-0.859	-0.799	-0.221
5	Fe I	4015.465	4.0434	1.407	7.44 ^(b)	33.3	0.610	-0.781	-0.775	-0.770	-0.787
6	Ni II	4015.474	3.9028	0.619	6.25 ^(a)	99.0	0.033	-2.419	-2.154	-2.529	-2.429
7	Nd II	4015.545	1.6766	0.886	1.50 ^(c)	93.9	0.026	0.080	0.365	0.118	0.589
8	Fe I	4015.605	4.0434	1.394	7.44 ^(b)	4.9	0.849	-0.515	-0.429	-0.513	-0.542
9	Ce II	4015.875	1.0083	0.881	1.55 ^(c)	18.9	0.079	-0.087	0.078	-0.076	0.018
10	Fe I	4015.986	4.1235	2.000	7.44 ^(b)	2.4	0.248	-1.928	-1.781	-1.925	-1.922
11	Ni I	4016.068	3.9735	1.388	6.25 ^(a)	13.3	0.064	-1.870	-1.724	-1.876	-1.760
12	Ti I	4016.274	2.0658	1.762	4.99 ^(a)	3.6	0.140	-0.714	-0.802	-0.691	-0.733
13	Fe I	4016.419	3.1775	0.682	7.44 ^(b)	1.0	0.865	-1.600			-1.633
14	Fe I	4016.541	2.6399	1.197	7.44 ^(b)	3.7	0.213	-3.513	-3.414	-3.505	-3.498
15	Mn I	4016.658	4.1926	0.035	5.39 ^(a)	40.4	0.012	-1.160	-1.348	-1.192	-0.876
16	Co II	4016.685	3.0177	1.293	4.92 ^(a)	28.3	0.018	-2.905	-2.835	-2.851	-3.197
17	Fe I	4016.792	4.1195	2.002	7.44 ^(b)	63.2	0.057	-2.576	-2.723	-2.574	-2.357
18	Co I	4016.793	3.5124	1.213	4.92 ^(a)	38.2	0.100	-0.547	-0.459	-0.545	-0.792

Notes. The columns represent line number in the spectral region used for analysis (see Fig. 9), chemical element, central wavelength λ_0 , the energy of lower level E_i , effective Landé factor g_{eff} , element abundance A , blending factor b , line core depth D_{core} , $\log(gf)_0$ parameter and inferred $\log(gf)$ in mode 2 averaged over all pixels ($\log(gf)_{\text{mode2}}$), mode 3 ($\log(gf)_{\text{mode3}}$) and using the spatially average method ($\log(gf)_{\text{spat}}$). All atomic data are from Kurucz’s database, except for the Landé factor and abundances. The effective Landé factors are taken from the VALD database, except for line 10, which is computed assuming LS coupling.

References. ^(a)Grevesse & Anders (1991). ^(b)Asplund et al. (2000). ^(c)Anders & Grevesse (1989).

lines in this region whose parameters are given in Table 1. The blending factor for each line was estimated using the method in Laverick et al. (2017). This factor quantifies the extent of overlap between the core of a given absorption line and its neighbouring lines. The smaller the blending factor is, the smaller the amount of overlap.

The Stokes profiles were computed at the disc centre ($\mu = 1$) under the LTE approximation using `globin`. The computed spectra were normalized using the continuum intensity computed from the HSRA atmospheric model³ (Gingerich et al. 1971). The differences between the coupled method and the pixel-by-pixel method were tested under ideal conditions by disregarding observational effects such as stray light, the finite resolving power of spectrographs and noise.

In the inversion, we fitted for the height stratification of the atmospheric parameters: temperature, LOS velocity, and magnetic field vector. The temperature was inferred using four nodes placed at optical depth values of $\log \tau = (-2.5, -1.5, -0.5, 0.4)$ while the remaining atmospheric parameters were inferred at three nodes, $\log \tau = (-2.2, -1.1, 0)$. Here and in the rest of the paper, the optical depth scale is specified at the reference wavelength of 5000 Å. The initial values for the atmospheric parameters were uniform across all pixels.

Among the atomic parameters, we considered $\log(gf)$ and the central wavelength shift $\Delta\lambda$ of all lines as free parameters, except for line 13, whose parameters were kept fixed during inversions in both the pixel-to-pixel and coupled methods. Line 13 is the strongest and the least blended within the spectral range. Fixing the atomic parameters of one spectral line was necessary for the absolute wavelength and LOS velocity estimate. Additionally, it improved the inference reliability of atomic parameters for other lines in the spectral region.

³ Any other mean atmospheric model, such as FAL-C or HOLMUL, could be used to compute the continuum intensity.

The initial values for the atomic parameters were randomized with a Gaussian distribution centred at the exact value (i.e. the value used for computing the synthetic spectra from the simulations) with a standard deviation of 0.2 dex for $\log(gf)$ and 5 mÅ for the $\Delta\lambda$ parameter. This initialization is compatible with the expected uncertainties in these parameters and reflects the realistic situation we have to deal with when applying the method to observed spectra. We limited the allowed ranges for $\log(gf)$ and $\Delta\lambda$ to $[-2.0, 1.5]$ dex and $[-30, 22]$ mÅ with respect to the exact values. These limits are in the range of the expected uncertainties in the atomic databases and can be adjusted for each spectral line independently. The asymmetric boundaries were chosen to remove any possibility for a zero average value, which could occur from averaging inferred values from symmetric boundaries.

To compare results from the coupled method with the pixel-by-pixel method, we performed the inversions in three different modes:

- Mode 1: Only atmospheric parameters are inverted for each pixel individually. The atomic parameters were fixed to the values used to compute the reference spectra. This inversion allowed us to retrieve the best possible stratification of atmospheric parameters with the given node settings and initial parameter values since all atomic parameters were assumed to be accurately known.
- Mode 2: Both atomic and atmospheric parameters were inverted for each pixel independently (pixel-by-pixel method). Here, the inversion retrieved different atomic parameters for every spectral line and every pixel.
- Mode 3: Atmospheric parameters varied between pixels, while atomic parameters were inverted globally (i.e. the coupled method). Only one set of atomic parameters ($\log(gf)$ and $\Delta\lambda$) was obtained per spectral line.

Mode 1 inversion results were used as a reference to which the inversion results from mode 2 and 3 are compared. Any large

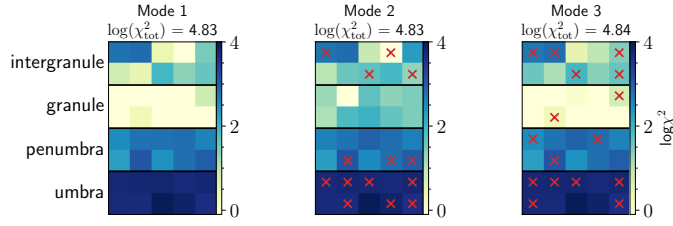


Fig. 2. χ^2 values for all three modes on a logarithmic scale. The total χ^2 value, the sum of the χ^2 values from all pixels, is displayed at the top of each panel. Pixels marked with a red cross in panels of mode 2 and 3 correspond to pixels with a lower χ^2 than the χ^2 value in mode 1.

deviations in the retrieved atmospheric parameters in mode 2 and 3, compared to the retrieved atmospheric parameters in mode 1, can likely be attributed to the poor atomic parameters in these two modes. All three modes were run with the same initial values of the atmospheric parameters, and the same initial atomic parameters in mode 2 and 3.

The χ^2 values⁴ for all selected atmospheres (referred to as pixels from now on) from all three modes are displayed in Fig. 2. We note that these pixels are not spatially connected (yellow dots in Fig. 1d). Inversion mode 1 shows a very low χ^2 value for pixels within the granules, whereas the umbral and penumbral pixels show higher values. The convergence to lower χ^2 in these pixels could be improved by changing the initial parameters or changing the convergence parameters of the LM algorithm. However, this does not guarantee that the χ^2 value in other pixels would also improve simultaneously.

The inversion mode 2 and 3 show a very good fit to the Stokes spectra that is comparable to mode 1, with slight differences in individual pixels. Those pixels in mode 2 and 3 with χ^2 value lower than in mode 1 are marked with red crosses in Fig. 2. With a larger number of free parameters, the inversion algorithm found a lower minimum in the χ^2 hypersurface.

The quality of the retrieved atmospheric parameters stratification in each pixel for all three inversion modes was quantified using the root-mean-square-deviation (RMSD) estimator. We defined the parameter's RMSD (P_{RMSD}) as the root-mean-square difference between the retrieved parameter stratification (P_{inv}) and the original parameter stratification in the MHD cube (P_{MHD}) at all depths between the highest and lowest nodes on the interpolated $\log \tau$ grid:

$$P_{\text{RMSD}} = \sqrt{\frac{1}{N_d} \sum_{i=1}^{N_d} (P_{\text{inv}_i} - P_{\text{MHD}_i})^2}, \quad (3)$$

where N_d is the number of fine grid points in the atmosphere between the lowest and the highest node of the parameter P . The extrapolated depth points above the highest and below the deepest node were excluded from the computation of the P_{RMSD} to focus on the region of the atmospheres where the lines are formed.

The P_{RMSD} values for the temperature, LOS velocity and magnetic field strength for each pixel are displayed in Fig. 3.

⁴ For direct comparison, χ^2 was calculated for each pixel as $\chi^2 = \frac{1}{N} \sum_{i=1}^N w_i^2 (\mathbf{O}_i - \mathbf{S}_i)^2$, where N is the number of wavelengths (running over each Stokes parameter), w_i is the weighting of each Stokes component and each wavelength, \mathbf{O} is a synthetic Stokes spectrum from selected pixels and \mathbf{S} is an inverted Stokes spectrum. Both \mathbf{O} and \mathbf{S} were normalized to the local Stokes I continuum value. The same weights were used in all three modes.

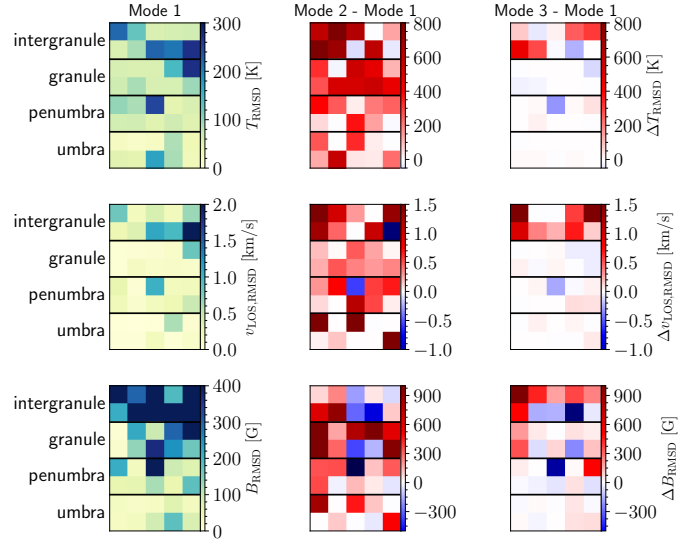


Fig. 3. Root-mean-square deviation of the retrieved atmospheric parameters' stratification calculated from the highest to the lowest nodes in each inversion mode for each pixel. The first column shows the P_{RMSD} for mode 1, while the other two columns display differences in P_{RMSD} from mode 2 and mode 3 from mode 1. The P_{RMSD} measures are given for temperature (first row), LOS velocity (second row) and the magnetic field strength (third row).

The first column shows the P_{RMSD} for mode 1 while the last two columns display the differences in P_{RMSD} for mode 2 and 3 relative to mode 1 (ΔP_{RMSD} value). Negative differences indicate a better retrieval of atmospheric parameters in mode 2 and 3 compared to the atmospheric parameters retrieved in mode 1. Overall, mode 3 shows a better retrieval of atmospheric parameters compared to mode 2, and the values retrieved are comparable to mode 1 inversion results.

A temperature offset was observed in some pixels in mode 2 that are related to umbra and penumbra features (see Fig. 3). In these pixels, the LM algorithm managed to achieve the same or lower χ^2 in comparison to the one obtained in mode 1. This is because, at every pixel in mode 2, the code has the freedom to adjust $\log(gf)$ and other atmospheric parameters independently to fit the line profile. When the algorithm chooses to over-tweak the $\log(gf)$ to get the desired fit, this can result in an offset of parameters like temperature that affects a line profile in a way similar to the $\log(gf)$. This cross talk between atomic and atmospheric parameters can result in large deviations in the retrieved values, while achieving a satisfactory fit to the Stokes profiles. One of the reasons for this was the use of a simple four-node representation for the temperature stratification in the inversions. It is possible that such a simple model was not adequate to accurately represent the complex height stratification found in the MHD atmospheres. On the other hand, the retrieved atmospheric parameters from mode 3 showed a smaller deviation from the values retrieved in mode 1, which indicates that the coupling between pixels weakens the cross talk between the temperature and the $\log(gf)$.

A comparison of the retrieved parameters stratification in all three modes to the MHD stratification for a granule atmosphere is displayed in Fig. 4. This figure exemplifies the highly complex stratifications typically found in MHD simulations. The corresponding synthetic Stokes spectra and the best fit for the three modes are displayed in Fig. 5, where mode 3 shows the best fit to the observed spectrum.

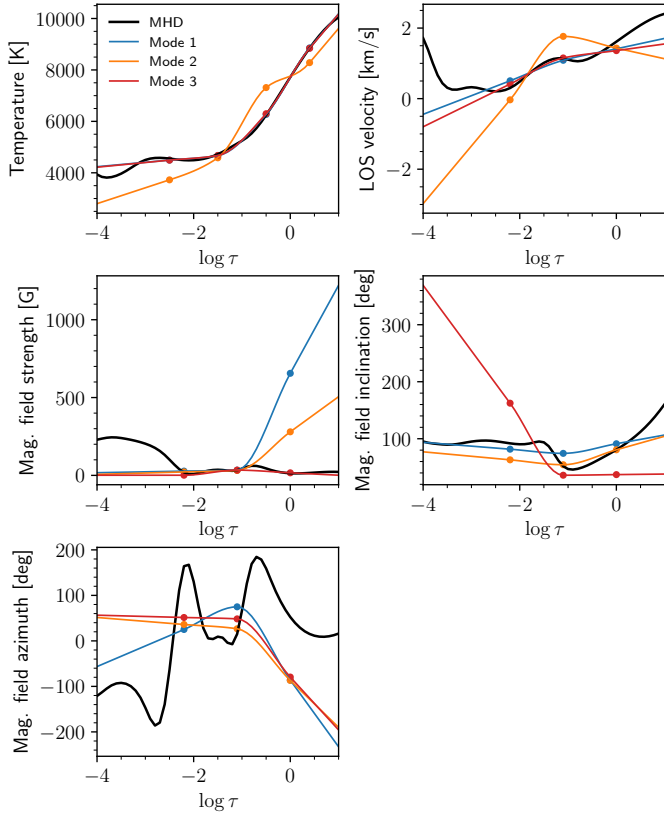


Fig. 4. Stratifications of various atmospheric parameters for a granule atmosphere. Curves with different colours represent the stratification of the original MHD simulation and different inversion modes (see legend in the top-left panel). Circles represent the node positions.

The comparison between the retrieved $\log(gf)_{\text{inv}}$ in mode 2 and 3 and the $\log(gf)_0$ is shown in Fig. 6. The mode 2 results represent the mean value from the $\log(gf)$ values of all considered pixels. This not only improved the statistical significance for the mode 2 results but also allowed for a fair comparison with the mode 3 results, which always takes into account all considered pixels. A preliminary analysis shows that no significant improvement in the results from mode 2 can be achieved by increasing the number of pixels. Also, it is to be expected that the optimal number of pixels required for each of these modes will depend on the number of lines for which the atomic parameters are to be determined and the amount of line blending. The influence of the chosen number of pixels on the results from both, mode 2 and 3, will be investigated in detail in a follow-up study.

As a measure of the quality of the $\log(gf)$ inference for mode 2, we used the mean and the standard error of the $\log(gf)_{\text{inv}} - \log(gf)_0$. The mode 2 inversion managed to retrieve the $\log(gf)$ values to an accuracy of 0.025 dex of the exact value, while mode 3 performs better, with an accuracy of only 0.004 dex. These values were computed by averaging the retrieved values weighted by their line core depth, giving larger weight to stronger and less blended lines. The standard error of $\Delta\lambda$ was below 1 mÅ for both modes.

Spectral line 4 has the largest standard error of $\log(gf)$ in mode 2 and the mean $\log(gf)$ close to the expected value (Fig. 6, see also Fig. 7). This is likely a result of the severe blending with line 3: changes in $\log(gf)$ of line 4 will not produce a significant signature in the spectrum, which hinders the inversion in retrieving a reliable value. A similar situation also holds for line 7,

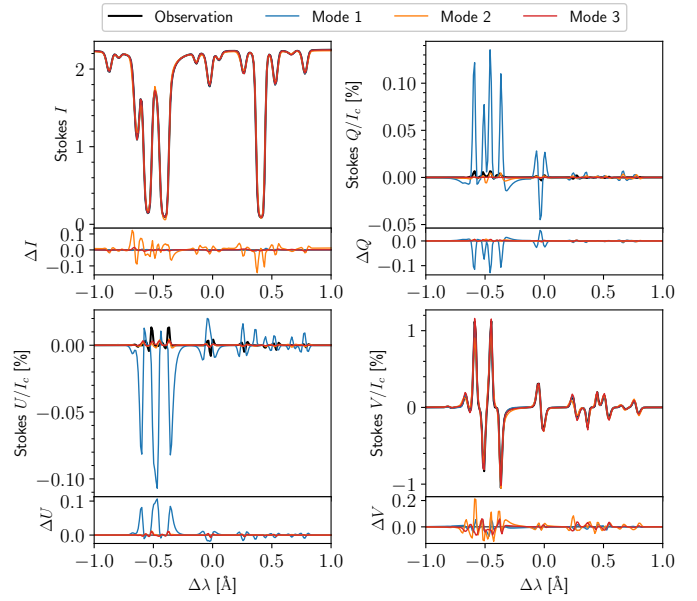


Fig. 5. Comparison of Stokes spectra in all inversion modes for a granule atmosphere whose atmospheric stratification is displayed in Fig. 4.

which forms a blend with line 8. Figure 7 shows that the uncertainty of the central wavelength of the spectral line can have a large impact on the retrieval of the exact $\log(gf)$ value. This is especially seen in line 4.

The problem in retrieving atomic parameters caused by the blending of spectral lines is illustrated by comparing the results for lines 15 and 16. Figure 6 shows that line 15 is weaker while line 16 is stronger than expected in mode 2. However, line 15 also shifts to shorter wavelengths, moving it further away from line 16. To reproduce the synthetic profile, the inversion code decreased the $\log(gf)$ of line 15 and increased it for line 16. This interplay between the $\log(gf)$ values of these lines is visible in Fig. 7. This problem was alleviated in mode 3, where no such large difference in retrieved $\log(gf)$ or $\Delta\lambda$ is observed. A similar explanation for the strongly blended pair of lines 17 and 18 also holds. This result clearly shows the benefits of using the coupled method to retrieve reliable atomic parameters for blended lines compared to the pixel-by-pixel method.

4. Test of the spatially averaged method

The standard method of inferring atomic parameters from observations relies on synthesizing a spectrum from a representative atmospheric model and adjusting only the atomic parameters until the best fit to a spatially averaged solar or stellar spectrum is achieved. We call this method the ‘spatially averaged method’. The information content in a single spectrum is lower than when using many spectra emerging from different atmospheric conditions. Therefore, the method works best when applied to strong and isolated spectral lines such as those found in the visible and the infrared spectral regions (e.g. Gurtovenko & Kostik 1981, 1982; Thévenin 1989, 1990; Borrero et al. 2003; Bigot & Thévenin 2006). When the lines are blended (i.e. their absorption profiles overlap in wavelength), the method requires simultaneous treatment of all blended spectral lines (Borrero et al. 2003) or a proper de-blending of the spectral line of interest (Shchukina & Vasil’eva 2013).

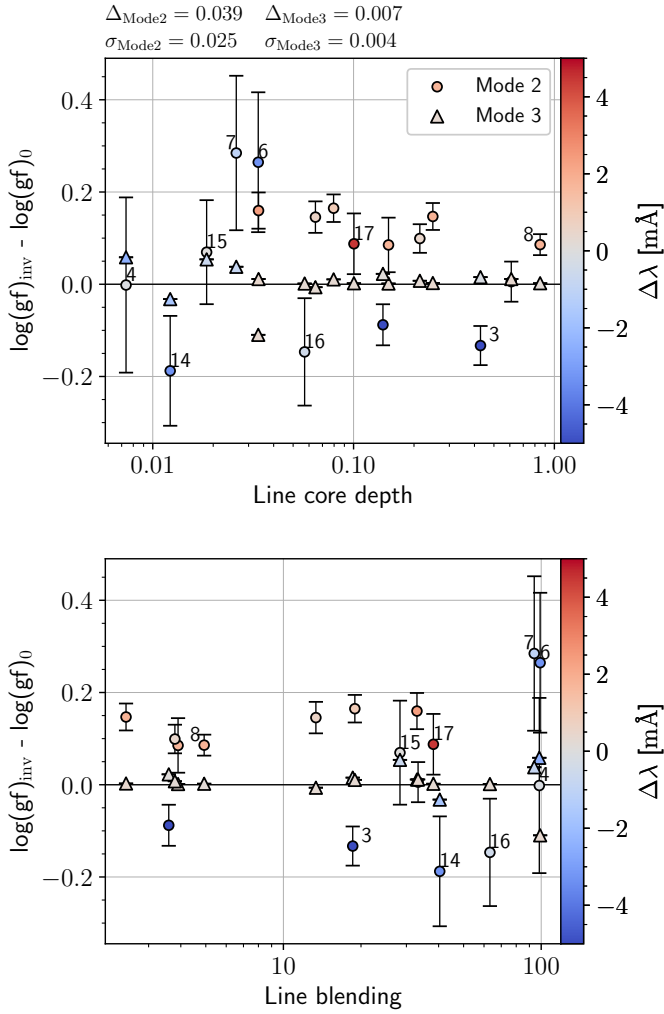


Fig. 6. Comparison of the atomic parameters, $\log(gf)$ and $\Delta\lambda$, for mode 2 and 3 inversions. Top panel: the y-axis shows the difference between inferred $\log(gf)_{\text{inv}}$ and exact $\log(gf)_0$ values where on the x-axis we plot the line core depth. The averaged values in mode 2 are marked in circles and the values from mode 3 in triangles. Markers are coloured based on the $\Delta\lambda$ parameter. The top of the panel lists mean Δ and standard error σ of $\log(gf)_{\text{inv}} - \log(gf)_0$ for each mode. The error bars are displayed only for mode 2, marking the 1σ level. Mode 3 retrieves a unique global value for each line, and we do not have any measure of error in this case. Bottom panel: the comparison of $\log(gf)_{\text{inv}} - \log(gf)_0$ with the line blending factor listed in Table 1. Markers shape and colour have the same meaning as in the top panel.

We performed a test of the spatially averaged method for the same spectral region as was used in Sect. 3, that is, 4015–4017 Å, containing 18 spectral lines (Table 1). Here, we computed a spatially averaged spectrum from another quiet-sun MURaM simulation having a mean magnetic field of strength of 50 G. We expect that this spatially averaged spectrum is rather similar to the observed average spectrum of the quiet Sun.

We opted to infer a representative 1D atmospheric model from this average spectrum, assuming that the atomic parameters are known and kept fixed during the inversion. A similar approach was also followed in Borrero et al. (2003). Additionally, the Landé factors of all synthesized lines were set to zero to remove the effect of the broadening introduced by the magnetic field. This simplified our inferred atmospheric model by eliminating the need to infer the magnetic field vector.

We used the `globin` code to infer the temperature T in four nodes at $\log \tau = (-2.2, -1.5, -0.8, 0)$, line-of-sight velocity v_{LOS} in two nodes at $\log \tau = (-2.2, 0)$, micro-turbulent velocity v_{mic} in three nodes at $\log \tau = (-2.2, -1.1, 0)$, and depth-independent macro-turbulent velocity v_{mac} . The v_{mac} and depth-dependent v_{mic} allowed us to properly model the broadening of spectral lines resulting from the spatial averaging of spectra from granules and intergranular lanes with different brightness and surface area coverage, and harbouring oppositely directed flows. The spatial averaging also produced asymmetric line profiles that were modelled by the height-dependent line-of-sight velocity stratification.

The inferred atmospheric parameters are displayed in Fig. 8 with $v_{\text{mac}} = 2.05 \text{ km s}^{-1}$, where the comparison of the spatially averaged and inverted spectra is displayed in Fig. 9. The retrieved micro-turbulent velocity is in accordance with what is generally found in any 1D atmospheric model, such as FAL-C (Fontenla et al. 1993), with higher values in the deepest layer that decrease with height in the photosphere to values below 1 km s^{-1} . The retrieved line-of-sight velocity shows a small gradient.

The comparison of spectra in Fig. 9 shows that the atmospheric model produces a good match in the cores of many lines, with some difficulties in reproducing the cores of very weak lines. The alteration of atmospheric parameters at heights to which the cores of these lines are sensitive does not significantly improve the χ^2 value. This example shows that it is impossible to simultaneously reproduce the entire profiles of different lines in a spatially averaged spectrum using a 1D atmospheric model. In a more rigorous examination, Uitenbroek & Criscuoli (2011) showed that 1D atmospheric models, constructed by averaging the individual atmospheres within a simulated 3D MHD atmosphere of the Sun, do not reproduce the average spectrum emerging from the 3D MHD atmospheric model very well. This result is a consequence of the radiative transfer equation being non-linear that leads to a non-linear mapping of atmospheric parameters onto the spectrum. Additionally, Uitenbroek & Criscuoli (2011) found that if an atmospheric model reproduces the spectrum in a specific wavelength range, it need not reproduce other wavelength ranges equally well.

We have experimented with different numbers of nodes in micro-turbulent velocity and temperature, but none of the models provided as good a match to the spatially averaged spectrum as the chosen one (see Figs. 8 and 9), which is in agreement with the findings by Uitenbroek & Criscuoli (2011). Therefore, with all the limitations of a 1D atmospheric model, we deemed that the spectrum from the inferred atmospheric model was the best possible representation of the spatially averaged spectrum from the quiet-sun MURaM cube (or at least very close to such a representation).

We then proceed to infer the atomic line parameters from the spatially averaged spectrum using the inferred atmospheric model. We assumed that the $\log(gf)$ parameters of all lines are unknown and need to be retrieved using the `globin` code. In this test, we assumed the central wavelengths of the lines to be known to ensure that there is no additional cross talk between any uncertainty in the wavelengths with the $\log(gf)$ parameter. The initial $\log(gf)$ values were randomized around the exact values, as was done in Sect. 3.

The inferred values for both the spatially averaged and the coupled method from Sect. 3 are displayed in Fig. 10. Both methods can reproduce the $\log(gf)$ and $\Delta\lambda$ values reasonably well for strong (line core depth ≥ 0.1) and unblended (line blending ≤ 10) spectral lines. But even for the stronger lines, the

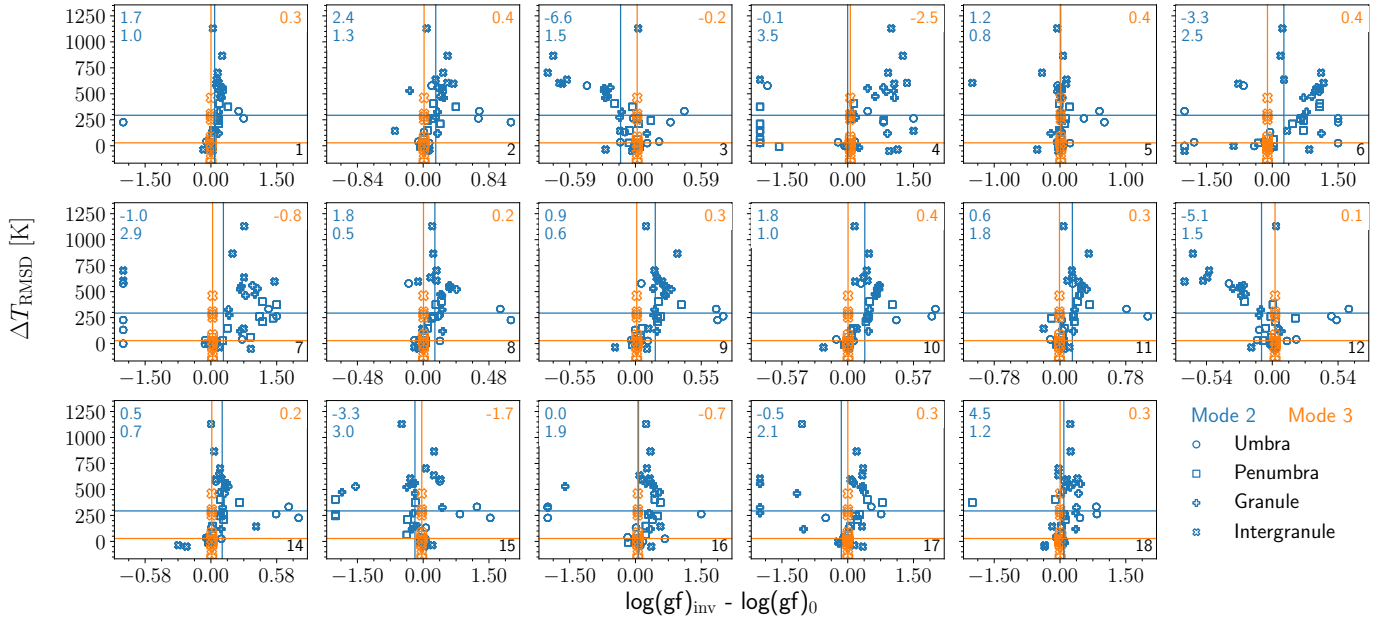


Fig. 7. Scatter plots of the $\log(gf)_{\text{inv}} - \log(gf)_0$ (x -axis) versus ΔT_{RMSD} values (y -axis) for mode 2 (in blue) and mode 3 (in orange). Each subplot corresponds to one spectral line with the line number from Table 1 given in the lower-right corner. The values from each pixel are represented with different symbols: circle (umbra), square (penumbra), plus (granule), and cross (intergranule). Horizontal and vertical lines mark the average of ΔT_{RMSD} and $\log(gf)_{\text{inv}} - \log(gf)_0$ in each inversion mode, respectively. The values in the upper left corner of each panel are the mean and standard error of $\Delta\lambda$ in mode 2, and in the upper right corner is the inferred $\Delta\lambda$ in mode 3 in units of mÅ. The standard error of $\Delta\lambda$ is zero in mode 3 since it retrieves a unique global value.

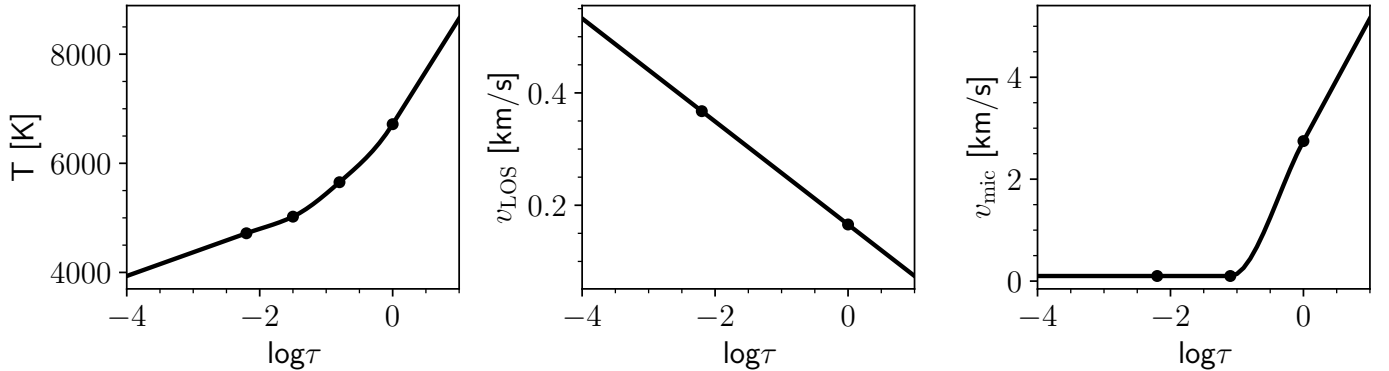


Fig. 8. Inferred atmospheric model (temperature, line-of-sight velocity, and micro-turbulent velocity) from the spatially averaged spectrum synthesized from the quiet-sun MURaM cube.

coupled method produced better results, visible in the improved match of the line cores of, for example, lines 8 and 13 (see Fig. 9). The retrieved $\log(gf)$ values for lines 8 and 13, listed in Table 1, show deviations that are comparable to errors expected for lines in the NUV (Shchukina & Vasil'eva 2013).

The results displayed in Fig. 10 also indicate that the spatially averaged method struggles to recover correct atomic parameters for weaker and blended lines. This is expected and demonstrates the limitation of the spatially averaged method, where the 1D atmospheric model cannot represent the true temperature stratification in the 3D MHD cube (Uitenbroek & Criscuoli 2011). The inversion code mitigates this problem by retrieving an incorrect value of $\log(gf)$ to fit the spatially averaged spectrum better.

It is interesting to note the results achieved for lines 3 and 5 that are rather strong lines but show considerable blending with the wings of line 8. The `globin` code retrieved sufficiently accurate $\log(gf)$ values for these two lines because of their measur-

able impact on the χ^2 value. The slight change in their $\log(gf)$ value results in a significant change to the quality of the fit.

In the presented case, line blending is most prominent in spectral lines 15 and 16 and 17 and 18, where the blending is so strong that both lines appear as a single spectral line. Weak and blended spectral lines showed the largest deviation of the retrieved $\log(gf)_{\text{inv}}$ from the expected $\log(gf)_0$ (Fig. 10), especially when they are weaker than the rest of the lines in their vicinity.

Comparison with the results from Sect. 3 shows the advantage of using the coupled instead of the spatially averaged method, especially for weak and blended spectral lines. Lacking the precise $\log(gf)$ values for weak lines can hinder reliable retrieval of atmospheric parameters from the strong lines with accurate $\log(gf)$ values. This is particularly relevant when applying the spatially averaged method to the NUV spectral lines where the line blending is more pronounced than in the visible and infrared regions.

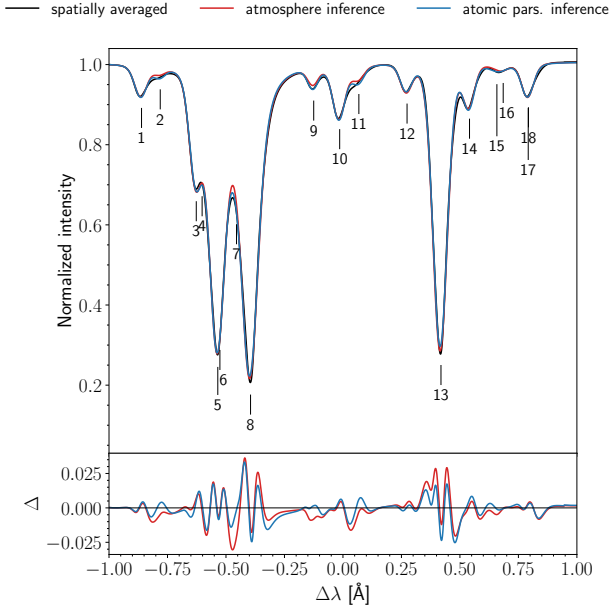


Fig. 9. Comparison between the spatially averaged spectrum from the quiet-sun MURaM cube (black line), the spectrum from the inference of the atmospheric model (red line), and the spectrum from the inference of the atomic line parameters (blue line). In the bottom panel, we show the difference between spectra compared to the spatially averaged one.

5. Conclusions

To infer the stratification of parameters of the solar atmosphere from spectropolarimetric observations, we require precise atomic data of spectral lines (excitation potential of energy levels, elemental abundance, transition probability $\log(gf)$, central wavelength, collisional broadening, etc.). The atomic parameters are also important in stellar physics for inferring the effective temperature, surface gravity and elemental abundances of different stars (Bigot & Thévenin 2006; Boeche & Grebel 2016; Heiter et al. 2021).

In this paper, we presented a new inversion method (named here as the coupled method) that retrieves atmospheric parameters at every spatial pixel, along with precise values of the atomic parameters, even for heavily blended spectral lines. Only a single set of the latter is determined, as they are not allowed to vary from one spatial position to another. This is achieved by requiring the Stokes spectra from all spatial pixels in the given field of view to be fit simultaneously, with the atmospheric parameters being free parameters individually for every pixel, and the atomic parameters being one 'global' set of values for all pixels. This method, implemented in the inversion code `globin`, follows a similar approach as the method of spatially coupled inversions by van Noort (2012).

We tested the coupled method and compared its results with those from the pixel-by-pixel method for the spectral region 4015 Å–4017 Å containing 18, mainly blended, spectral lines. We synthesized line profiles assuming the LTE approximation using the atmospheric models of the umbra, penumbra, granule and intergranular lanes extracted from the MURaM simulation of a sunspot. The coupled method is able to retrieve atmospheric parameters, $\log(gf)$ and $\Delta\lambda$, with high precision, whereas the pixel-by-pixel method shows a large deviation in the temperature and $\log(gf)$ from the reference values. This deviation is due to the blending of spectral lines and, to some extent, due to the cross talk between inversion parameters (mainly between tem-

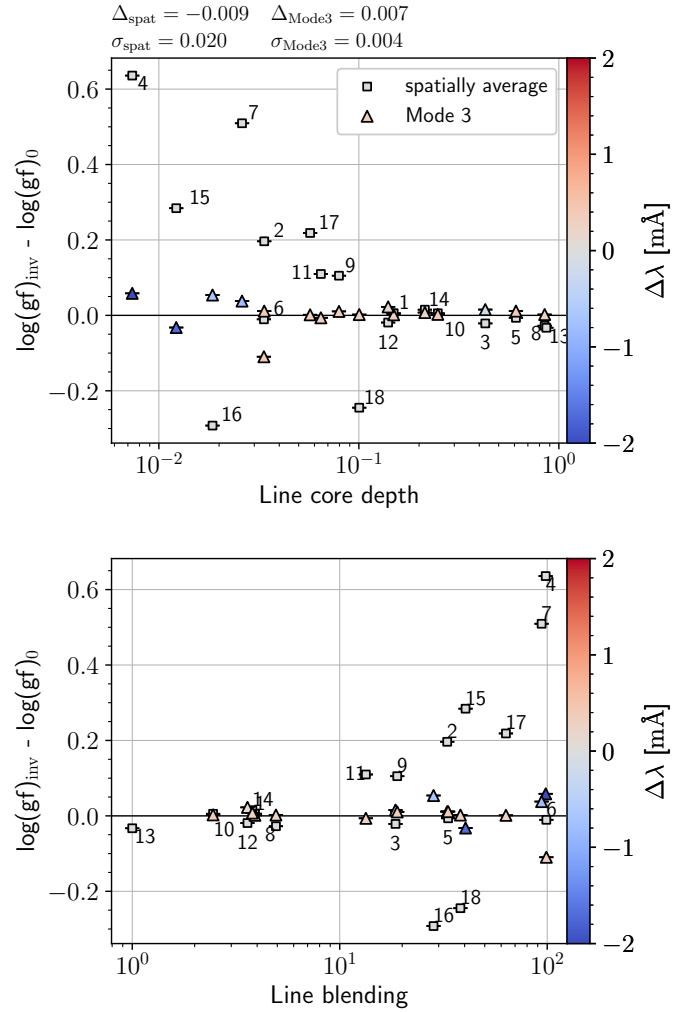


Fig. 10. Comparison of the inferred atomic line parameters using the spatially average and the coupled methods. Top panel: the difference between inferred and expected $\log(gf)$ values versus the line core depth from the spatially averaged method (squares) and from mode 3 in Sect. 3 (triangles). The colours of the symbols correspond to the wavelength shift $\Delta\lambda$ with respect to the central wavelength from Table 1 (only for mode 3 results). Bottom panel: the difference between inferred and expected $\log(gf)$ values versus the line blending factor from Table 1. Colour coding is the same as in the top panel.

perature and $\log(gf)$ and $\log(gf)$ and $\Delta\lambda$). The pixel-by-pixel method is unable to decouple line contributions to blended profiles, thus inferring erroneous parameter values.

The inverted line profiles are affected by the goodness of the atmospheric parameterization used to reproduce the synthetic line profiles. The simple approximation of using a few nodes might not be adequate to represent the highly complex stratification found in MHD atmospheric models of different solar features. As a result of this, we have always contrasted our results for atmospheric parameters to those retrieved in the inversion with fixed atomic parameters (equal to those used to synthesize observed spectra). We found that imposing a coupling in atomic parameters reduces the cross talk between atmospheric and atomic parameters, retrieving the stratification in atmospheric parameters as best as the chosen node parameterization can achieve from the given setup.

In producing the synthetic observations, we have ignored any instrumental and observational degradation of spectra, allowing

us to compare the performances of the pixel-by-pixel and coupled methods under ideal conditions. Introducing the degradation caused by the photon noise would mask very weak lines, causing the optimization algorithm to be insensitive to changes in the atomic parameters of those lines. This presents limitations on which lines we can apply our method to infer reliably the atomic parameters.

We showed how the coupled method can aid us in separating contributions of blended line profiles and retrieving accurate $\log(gf)$ values. This result is significant for the analysis of many-line spectropolarimetric observations, such as the NUV spectra expected from the SUSI instrument of the SUNRISE III observatory. The high density of spectral lines in this spectral region results in line blending. Additionally, there are considerable uncertainties in the knowledge of the atomic parameters for many of them. Besides the NUV lines, the coupled method can be used to retrieve the atomic parameters for lines in the visible and infrared wavelengths.

Another issue of analysing the NUV spectra and inferring the atomic parameters of those lines is properly accounting for the continuum opacity sources. In our study, we have ignored the impact of the continuum opacity contribution from unresolved lines that produce line haze (see e.g. Greve & Zwaan 1980; Rutten 2019) and assumed that we can model the true continuum. Not being able to reproduce the true continuum in the observed spectrum will be compensated by adjusting the retrieved temperature stratification and $\log(gf)$ of lines. Currently, the opacity fudge coefficients (e.g. Bruls et al. 1992) are used to enhance the calculated continuum opacity and reproduce the observed continuum.

Furthermore, we inferred the $\log(gf)$ values for the spectral lines in the same spectral range by fitting the spatially averaged synthetic spectrum computed from the quiet-sun MURaM atmospheric model. The 1D atmospheric model is inferred from the spatially averaged spectrum, assuming that atomic parameters are known. We then used this atmospheric model and repeated the inversion to infer only the $\log(gf)$ value of each line. The initial $\log(gf)$ values were displaced from the exact values by Gaussian randomization. The inferred atomic parameters for strong and unblended lines matched the exact values very well, while significant scattering was obtained for weaker and more blended spectral lines. This finding is in line with the use of the spatially averaged spectrum of the Sun to derive the $\log(gf)$ values assuming a representative 1D atmospheric model (the spatially averaged method). The spatially averaged method is generally restricted to isolated and strong spectral lines in the visible and infrared wavelengths (see for example Borrero et al. 2003).

In all the inversions presented in this paper, we assumed that the used line list is complete (all lines are known and accounted for during synthesis and inversion). This is not always the case for observed spectra. Disregarding one of the lines will result in an erroneous $\log(gf)$ value of other lines sufficiently blended with the disregarded line. The impact of missing lines on the inferred atmospheric stratification and atomic parameters is currently under investigation and will be reported in a separate publication. Additionally, the accuracy of the inferred $\log(gf)$ is also impacted by the poor knowledge of other atomic line parameters, such as the line central wavelength and the excitation potential of the lower energy level. We still rely heavily on laboratory measurements to provide precise values of these parameters. The poor knowledge of atomic parameters does not refer only to the line of interest, but also to other lines that are part of the blend.

We assumed in all inversions that lines are properly modelled in the LTE approximation, thus removing any discrepancies

between the synthetic and inverted line profiles that might arise from different modelling assumptions. NLTE effects, important for the formation of many spectral lines in the considered spectral region, can be modelled with `globin` and pose no additional problem in the inference of the above-mentioned atomic parameters. It is not yet clear how big of an impact the LTE approximation has on the inference of the $\log(gf)$ parameter for lines that should be treated in NLTE. We plan to investigate this for the widely used line pair of Fe I at 6301.5 Å and 6302.5 Å, which is shown to be sensitive to NLTE effects (Smitha et al. 2020).

We intend to extend the coupled method and add elemental abundance as another inversion parameter that is considered as a global parameter having a unique value across the observed field of view. Inverting a spectral region that contains many spectral lines of a single atomic element will impose additional constraints on the inference of the elemental abundance. We believe it is possible to disentangle the contribution of the elemental abundance and $\log(gf)$ to the line opacity by having at least one line in the observed spectral region with reliable $\log(gf)$. This will bring a completely new and independent way of determining the solar elemental abundances, which are currently determined mainly using 3D hydrodynamical models of the solar atmosphere (see for e.g. Asplund et al. 2009) that already use prescribed elemental abundances.

Acknowledgements. D.V. is funded by International Max Planck Research School (IMPRS) for Solar System Science at the University of Göttingen. We thank M. Rempel for kindly providing the MHD cubes. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101097844 – project WINSUN). This work has made use of the VALD database, operated at Uppsala University, the Institute of Astronomy RAS in Moscow, and the University of Vienna.

References

- Anders, E., & Grevesse, N. 1989, *Geochim. Cosmochim. Acta*, **53**, 197
- Asplund, M., Nordlund, Å., Trampedach, R., & Stein, R. F. 2000, *A&A*, **359**, 743
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Bigot, L., & Thévenin, F. 2006, *MNRAS*, **372**, 609
- Blackwell, D. E., & Collins, B. S. 1972, *MNRAS*, **157**, 255
- Boeche, C., & Grebel, E. K. 2016, *A&A*, **587**, A2
- Borrero, J. M., Bellot Rubio, L. R., Barklem, P. S., & del Toro Iniesta, J. C. 2003, *A&A*, **404**, 749
- Bruls, J. H. M. J., Rutten, R. J., & Shchukina, N. G. 1992, *A&A*, **265**, 237
- Castellanos Durán, J. S., Korpi-Lagg, A., & Solanki, S. K. 2023, *ApJ*, **952**, 162
- de la Cruz Rodríguez, J. 2019, *A&A*, **631**, A153
- de la Cruz Rodríguez, J., & Piskunov, N. 2013, *ApJ*, **764**, 33
- de la Cruz Rodríguez, J., & van Noort, M. 2017, *Space Sci. Rev.*, **210**, 109
- de la Cruz Rodríguez, J., Leenaarts, J., Danilovic, S., & Uitenbroek, H. 2019, *A&A*, **623**, A74
- del Toro Iniesta, J. C. 2003, *Introduction to Spectropolarimetry* (Cambridge University Press)
- del Toro Iniesta, J. C., & Ruiz Cobo, B. 2016, *Liv. Rev. Sol. Phys.*, **13**, 4
- Den Hartog, E. A., Lawler, J. E., & Smeden, C. 2005, *Phys. Scr. Vol. T*, **119**, 45
- Feller, A., Gandorfer, A., Iglesias, F. A., et al. 2020, *SPIE Conf. Ser.*, **11447**, 11447AK
- Fontenla, J. M., Avrett, E. H., & Loeser, R. 1993, *ApJ*, **406**, 319
- Frutiger, C., Solanki, S. K., Fligge, M., & Bruls, J. H. M. J. 2000, *A&A*, **358**, 1109
- Gaia Collaboration (Prusti, T., et al.) 2016, *A&A*, **595**, A1
- Gingerich, O., Noyes, R. W., Kalkofen, W., & Cuny, Y. 1971, *Sol. Phys.*, **18**, 347
- Greve, A., & Zwaan, C. 1980, *A&A*, **90**, 239
- Grevesse, N., & Anders, E. 1991, *Solar Interior and Atmosphere* (Tucson, AZ: University of Arizona Press), 1227
- Gurtovenko, E. A., & Kostik, R. I. 1981, *A&AS*, **46**, 239
- Gurtovenko, E. A., & Kostik, R. I. 1982, *A&AS*, **47**, 193
- Heiter, U., Lind, K., Bergemann, M., et al. 2021, *A&A*, **645**, A106
- Kramida, A., Ralchenko, Yu., Reader, J., & NIST ASD Team 2022, *NIST Atomic Spectra Database (ver. 5.10)* (Gaithersburg, MD: National Institute of Standards and Technology), <https://physics.nist.gov/asd>

- Kurucz, R., & Bell, B. 1995, *Atomic LineData*, eds. R. Kurucz, & B. Bell (Cambridge, Mass: Smithsonian Astrophysical Observatory), 23
- Laverick, M. 2019, PhD Thesis, KU Leuven, Belgium
- Laverick, M., Lobel, A., Royer, P., Martayan, C., & Merle, T. 2017, *Can. J. Phys.*, **95**, 843
- Levenberg, K. 1944, *Quart. Appl. Math.*, **2**, 164
- Marquardt, D. W. 1963, *J. Soc. Indust. Appl. Math.*, **11**, 431
- Martins, L. P., Coelho, P., Caproni, A., & Vitoriano, R. 2014, *MNRAS*, **442**, 1294
- Mihalas, D. 1978, *Stellar Atmospheres* (San Francisco: W.H. Freeman)
- Milić, I., & van Noort, M. 2018, *A&A*, **617**, A24
- Nave, G., Johansson, S., Learner, R. C. M., Thorne, A. P., & Brault, J. W. 1994, *ApJS*, **94**, 221
- Nave, G., Sansonetti, C. J., Townley-Smith, K., et al. 2017, *Can. J. Phys.*, **95**, 811
- Neckel, H., & Labs, D. 1984, *Sol. Phys.*, **90**, 205
- Orozco Suárez, D., Bellot Rubio, L. R., & del Toro Iniesta, J. C. 2007, *ApJ*, **662**, L31
- Peterson, R. C., & Kurucz, R. L. 2022, *ApJS*, **260**, 28
- Piskunov, N. E., Kupka, F., Ryabchikova, T. A., Weiss, W. W., & Jeffery, C. S. 1995, *A&AS*, **112**, 525
- Pradhan, A. K., & Saraph, H. E. 1977, *J. Phys. B At. Mol. Phys.*, **10**, 3365
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3rd edn. (Cambridge University Press)
- Quintero Noda, C., Shimizu, T., de la Cruz Rodríguez, J., et al. 2016, *MNRAS*, **459**, 3363
- Rempel, M. 2012, *ApJ*, **750**, 62
- Riethmüller, T. L., & Solanki, S. K. 2019, *A&A*, **622**, A36
- Ruiz Cobo, B., & del Toro Iniesta, J. C. 1992, *ApJ*, **398**, 375
- Rutten, R. J. 2019, *Sol. Phys.*, **294**, 165
- Rybicki, G. B., & Hummer, D. G. 1991, *A&A*, **245**, 171
- Shchukina, N. G., & Vasil'eva, I. E. 2013, *Kinematics Phys. Celestial Bodies*, **29**, 53
- Smitha, H. N., Holzreuter, R., van Noort, M., & Solanki, S. K. 2020, *A&A*, **633**, A157
- Socas-Navarro, H., de la Cruz Rodríguez, J., Asensio Ramos, A., Trujillo Bueno, J., & Ruiz Cobo, B. 2015, *A&A*, **577**, A7
- Solanki, S. K. 1987, PhD Thesis, Max-Planck-Institute for Solar System Research, Lindau, Germany
- Solanki, S. K., Barthol, P., Danilovic, S., et al. 2010, *ApJ*, **723**, L127
- Solanki, S. K., Riethmüller, T. L., Barthol, P., et al. 2017, *ApJS*, **229**, 2
- Thévenin, F. 1989, *A&AS*, **77**, 137
- Thévenin, F. 1990, *A&AS*, **82**, 179
- Tiwari, S. K., van Noort, M., Lagg, A., & Solanki, S. K. 2013, *A&A*, **557**, A25
- Tiwari, S. K., van Noort, M., Solanki, S. K., & Lagg, A. 2015, *A&A*, **583**, A119
- Trelles Arjona, J. C., Ruiz Cobo, B., & Martínez González, M. J. 2021, *A&A*, **648**, A68
- Uitenbroek, H. 2001, *ApJ*, **557**, 389
- Uitenbroek, H., & Criscuoli, S. 2011, *ApJ*, **736**, 69
- van Noort, M. 2012, *A&A*, **548**, A5
- van Noort, M., Lagg, A., Tiwari, S. K., & Solanki, S. K. 2013, *A&A*, **557**, A24
- Vögler, A., Shelyag, S., Schüssler, M., et al. 2005, *A&A*, **429**, 335

Appendix A: Pixel-by-pixel method

The inversion of spectropolarimetric observations is a non-linear optimization problem. A regularly used optimization scheme is the Levenberg-Marquardt algorithm (Levenberg 1944; Marquardt 1963, henceforth LM), which is a combination of the gradient descent and the Gauss-Newton methods for minimizing a merit function.

An important part of every optimization procedure is to define a merit function to be minimized through the iterative correction of free parameters. Spectropolarimetric inversion algorithms generally use χ^2 as a merit function, which is defined as (del Toro Iniesta 2003):

$$\chi^2(\mathbf{p}) = \frac{1}{N-n} \sum_{i=1}^N \frac{w_i^2}{\sigma_i^2} \cdot (\mathbf{O}_i - \mathbf{S}(\mathbf{p})_i)^2, \quad (\text{A.1})$$

where the index i goes over each wavelength point for all four Stokes components ($N = 4N_\lambda$, N_λ being the number of wavelength points), and n the number of free parameters. The factor w_i is the weight given to each wavelength, and σ_i is the corresponding noise at this wavelength. \mathbf{O}_i denotes the observed Stokes vector and $\mathbf{S}(\mathbf{p})_i$ denotes the synthetic Stokes vector, with \mathbf{p} representing a vector of free parameters (temperature T , line-of-sight velocity v_{LOS} , micro-turbulent velocity v_{mic} , magnetic field strength B , magnetic field inclination θ , magnetic field azimuth ϕ , $\log(gf)$, $\Delta\lambda$). The Stokes vector is defined as (I, Q, U, V) .

The LM algorithm uses the first derivative of a model function to minimize the merit function that defines a hypersurface in the inversion parameter space. The χ^2 hypersurface may have many local minima and usually only one global minimum corresponding to the best inversion parameters set. However, some local minima can be almost as deep as the global minimum because of the cross talk between different inversion parameters. The LM algorithm efficiently finds a minimum of the merit hypersurface but does not guarantee that the found minimum corresponds to the global one. By adjusting the weights w_i of different wavelengths in the spectrum, it is possible to change the shape of the merit hypersurface. Such an adjustment can produce a more pronounced global minimum and increase the efficiency of finding the global minimum of the merit function. Finding the best fit to the observed spectrum depends on how close the initial solution is to the optimal one, the complexity of the model function and the employed weights.

In the LM algorithm, an initial guess for the inversion parameters is assumed to be close to the global minimum. Expanding the merit function around the global minimum in parameter space using a second-order Taylor polynomial yields (Press et al. 2007):

$$\chi^2(\mathbf{p}) = \chi^2(\mathbf{p}_j) + \Delta\mathbf{p}_j^T \cdot \nabla\chi^2(\mathbf{p}_j) + \frac{1}{2}\Delta\mathbf{p}_j^T \cdot \mathcal{H} \cdot \Delta\mathbf{p}_j, \quad (\text{A.2})$$

where $\Delta\mathbf{p}_j = \mathbf{p} - \mathbf{p}_j$ is a correction for the parameter vector \mathbf{p}_j in the j -th iteration, \mathcal{H} is the Hessian matrix of the system, and T indicates the matrix transpose. Based on our assumption being at the minimum, we expect the gradient of χ^2 to be zero, $\nabla\chi^2(\mathbf{p}) = 0$. Thus, taking a gradient of Eq. (A.2), we obtain:

$$0 = \nabla\chi^2(\mathbf{p}_j) + \mathcal{H} \cdot \Delta\mathbf{p}_j, \quad (\text{A.3})$$

where $\frac{d\Delta\mathbf{p}_j}{d\mathbf{p}} = 1$. In this derivation, we used the identity:

$$\frac{d(\Delta\mathbf{p}_j^T \cdot \mathcal{H} \cdot \Delta\mathbf{p}_j)}{d\mathbf{p}} = (\mathcal{H} + \mathcal{H}^T) \cdot \Delta\mathbf{p}_j = 2\mathcal{H} \cdot \Delta\mathbf{p}_j,$$

where we applied the symmetry property of the Hessian matrix, $\mathcal{H} \equiv \mathcal{H}^T$. Solving Eq. (A.3) gives the correction $\Delta\mathbf{p}_j$ for the initial parameter vector that minimizes χ^2 . Here, we implicitly assume that the correction of the inversion parameters produces a linear change in the model, which produces a quadratic change in the χ^2 . Precisely this linearization of the non-linear model requires an iterative correction of inversion parameters.

The gradient of the merit function from Eq. (A.3) is:

$$-\frac{\partial\chi^2}{\partial p_k} = \frac{2}{N-n} \sum_{i=1}^N \frac{w_i^2}{\sigma_i^2} \cdot (\mathbf{O}_i - \mathbf{S}_i(\mathbf{p})) \cdot \frac{\partial\mathbf{S}_i}{\partial p_k}, \quad (\text{A.4})$$

where the term $\frac{d\mathbf{S}_i}{dp_k}$ is the response function of the Stokes spectrum to the k -th inversion parameter at wavelength i (del Toro Iniesta 2003). With further differentiation of Eq. (A.4) with respect to parameter p_l , we obtain the Hessian matrix elements:

$$\mathcal{H}_{l,k} = \frac{\partial^2\chi^2}{\partial p_l \partial p_k} = \frac{2}{N-n} \sum_{i=1}^N \frac{w_i^2}{\sigma_i^2} \cdot \left[\frac{\partial\mathbf{S}_i}{\partial p_k} \cdot \frac{\partial\mathbf{S}_i}{\partial p_l} - (\mathbf{O}_i - \mathbf{S}_i(\mathbf{p})) \cdot \frac{\partial^2\mathbf{S}_i}{\partial p_l \partial p_k} \right]. \quad (\text{A.5})$$

The second term under square brackets in Eq. (A.5) is usually disregarded in LM inversions (del Toro Iniesta 2003; de la Cruz Rodríguez et al. 2019). This is reasonable since we assumed the initial solution is close to the minimum. Therefore, $\mathbf{O} - \mathbf{S}(\mathbf{p})$ should be zero and the second-order derivatives will not influence the proposed parameter steps. Disregarding second-order derivatives ensures that the Hessian matrix is positive definite, and the step in parameter space leads to the minimization of χ^2 . In some occasions, the second-order derivatives can even lead to corrupted behaviour (Press et al. 2007).

Substituting Eqs. (A.4) and (A.5) into Eq. (A.3), and identifying new terms, we obtain:

$$\mathcal{H} \cdot \Delta \mathbf{p}_j = \mathcal{J}^T \cdot \Delta, \quad (\text{A.6})$$

where $\Delta_i = \sqrt{\frac{2}{N-n}} \cdot \frac{w_i}{\sigma_i} \cdot (\mathbf{O}_i - \mathbf{S}_i(\mathbf{p}_j))$ and \mathcal{J} is the Jacobian matrix of the system whose elements are given as:

$$\mathcal{J}_{i,k} = \sqrt{\frac{2}{N-n}} \frac{w_i}{\sigma_i} \frac{\partial \mathbf{S}_i}{\partial p_k}.$$

The linearization of the Hessian matrix allows us to express it through the Jacobian matrix as $\mathcal{H} = \mathcal{J}^T \mathcal{J}$.

Fitting non-linear models to observed data can lead to poor parameter corrections. To address this issue, the diagonal elements of the Hessian matrix are multiplied with a factor λ_M and added to the Hessian, yielding $\mathcal{H} = \mathcal{J}^T \mathcal{J} + \lambda_M \cdot \text{diag}(\mathcal{J}^T \mathcal{J})$. This factor is known as the Marquardt parameter and regulates the magnitude of the parameters correction (i.e. $\Delta p_k \propto 1/\lambda_M$). For fast convergence, the Marquardt parameter should be small enough (e.g. $\lambda_M < 10^{-2}$) but not too small to overstep the global minimum.

Appendix B: The coupled method – Algorithm layout

In the pixel-by-pixel method, the Jacobian and the Hessian matrices are constructed for every pixel where Eq. A.6 provides a parameter correction independently from other pixels. In the coupled method, a global parameter has a uniform value in all pixels whose retrieval requires a coupling between the pixels. This coupling is introduced by requiring a simultaneous match in the observed and inverted Stokes spectra from all the pixels. In the coupled method, inversion parameters are obtained by minimizing the sum of χ^2 from all the pixels:

$$\chi_{\text{global}}^2(\mathbf{p}) = \frac{1}{N - N_p} \sum_{a=1}^{N_{\text{atm}}} \sum_{i=1}^N \frac{w_i^2}{\sigma_i^2} \cdot (\mathbf{O}_{i,a} - \mathbf{S}(\mathbf{p})_{i,a})^2, \quad (\text{B.1})$$

where the index a goes over every atmosphere in the considered field-of-view containing the total number of atmospheres N_{atm} , and N_p is the total number of free parameters summed over all pixels. Here, the parameter vector \mathbf{p} contains parameters from every pixel in the field-of-view $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{N_{\text{atm}}})$.

In the coupled method, the functional dependence of the χ^2 function does not change except for the addition of a summation which runs over all pixels (atmospheres). Therefore, following the same procedure as in the pixel-by-pixel inversion, we derived an equation for the correction of the parameters for the coupled inversions:

$$\mathcal{H}_{\text{global}} \Delta \mathbf{p}_j = \mathcal{J}_{\text{global}}^T (\mathbf{O} - \mathbf{S}(\mathbf{p}_j)), \quad (\text{B.2})$$

where $\mathcal{H}_{\text{global}}$ and $\mathcal{J}_{\text{global}}$ are the global Hessian and Jacobian matrices of the system, respectively. The global Hessian and Jacobian matrices are connected in the same manner as in the pixel-by-pixel method. Therefore, to explain how the coupled method works, it is sufficient to derive the global Jacobian matrix.

In the pixel-by-pixel method, column k in the Jacobian matrix contains the response function of the Stokes vector \mathbf{R}_k to the k -th inversion parameter, where \mathbf{R}_k is a vector of length $4 \times N_\lambda$. In case of n inversion parameters, the Jacobian matrix for a single pixel has a dimension $(4 \times N_\lambda, n)$ and is written as $\mathcal{J} = (\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_n)$.

The global Jacobian matrix of the system is constructed by placing all the single-pixel Jacobian matrices on the diagonal:

$$\mathcal{J}_{\text{global}} = \begin{pmatrix} \mathcal{J}_1 & 0 & \cdot & \cdot & \cdot & & 0 \\ 0 & \mathcal{J}_2 & & & & & \cdot \\ \cdot & & \cdot & & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ 0 & & & & & \mathcal{J}_{N_{\text{atm}}-1} & 0 \\ & & & & & 0 & \mathcal{J}_{N_{\text{atm}}} \end{pmatrix}. \quad (\text{B.3})$$

This matrix is block-diagonal and ensures an uncoupled inversion of parameters for each pixel. The transposed form of a block-diagonal matrix is a block-diagonal matrix of transposed sub-matrices, resulting in a block-diagonal global Hessian matrix. The global Jacobian matrix corresponds to the system with $n \times N_{\text{atm}}$ number of free parameters and has a dimension $(4 \times N_\lambda \times N_{\text{atm}}, n \times N_{\text{atm}})$.

The block-diagonal form of the global Jacobian matrix has to be disrupted to achieve a coupled inversion. Therefore, let us assume that out of n parameters for each pixel we have l local (atmospheric) and g global (atomic) parameters ($n = l + g$). The atmospheric parameters vary between pixels due to the different physical structures of each pixel, while atomic parameters are the same for every pixel. In the coupled method, the total number of free parameters is therefore $N_{\text{atm}} \times l + g$, whereas in the pixel-by-pixel method this number is $N_{\text{atm}} \times (l + g)$. Further on, we assume that the inversion parameters in the vector \mathbf{p} are ordered from atmospheric to atomic.

Substituting the Jacobian sub-matrices with the response functions, we get:

$$\mathcal{J}_{\text{global}} = \begin{pmatrix} \mathbf{R}_1^1 \dots \mathbf{R}_l^1 & \mathbf{R}_{l+1}^1 \dots \mathbf{R}_{l+g}^1 & 0 & \dots & 0 \\ 0 & \mathbf{R}_1^2 \dots \mathbf{R}_l^2 & \mathbf{R}_{l+1}^2 \dots \mathbf{R}_{l+g}^2 & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \dots & \dots & \mathbf{R}_1^{N_{\text{atm}}} \dots \mathbf{R}_l^{N_{\text{atm}}} & \mathbf{R}_{l+1}^{N_{\text{atm}}} \dots \mathbf{R}_{l+g}^{N_{\text{atm}}} \end{pmatrix}, \quad (\text{B.4})$$

where the upper index in the response function corresponds to the pixel number.

To have the global inversion of atomic parameters, we have to reorder the response functions in this matrix. For a single global parameter, we take its response functions for all pixels and form a single column. Consequently, for the k -th parameter we have $(\mathbf{R}_k^1, \mathbf{R}_k^2, \dots, \mathbf{R}_k^{N_{\text{atm}}})^T$. Further, we repeat this process for each global parameter we have in the inversion. This results in a total of g columns that contain only the response functions to the global parameters. Then, we move these columns to the right side of the matrix. The rest of the matrix now contains the response functions to the local parameters only. These response functions form a sub-matrix that has a block-diagonal form. Therefore, the global Jacobian matrix after the reordering in the response functions is:

$$\mathcal{J}_{\text{global}} = \begin{pmatrix} \mathbf{R}_1^1 \dots \mathbf{R}_l^1 & 0 & \dots & 0 & \mathbf{R}_{l+1}^1 \dots \mathbf{R}_{l+g}^1 \\ 0 & \mathbf{R}_1^2 \dots \mathbf{R}_l^2 & \dots & \dots & \mathbf{R}_{l+1}^2 \dots \mathbf{R}_{l+g}^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & \vdots \\ 0 & \dots & \dots & \mathbf{R}_1^{N_{\text{atm}}} \dots \mathbf{R}_l^{N_{\text{atm}}} & \mathbf{R}_{l+1}^{N_{\text{atm}}} \dots \mathbf{R}_{l+g}^{N_{\text{atm}}} \end{pmatrix}. \quad (\text{B.5})$$

The block-diagonal form of the global Jacobian matrix is retained only for the local parameters and the coupling is introduced only for the inversion of global parameters. Corrections of global parameters are determined from spectrum differences in all pixels, while corrections for local parameters are determined from the spectral difference in a given pixel. This reordering keeps the pixel-by-pixel inversion of atmospheric parameters and introduces spatial coupling of, for example, atomic parameters.

The global Jacobian matrix with coupling in atomic parameters has dimension $(4 \times N_\lambda \times N_{\text{atm}}, N_{\text{atm}} \times l + g)$ which is lower in comparison to a global Jacobian matrix for uncoupled inversion due to the grouping of the response functions of the global parameters. This coupling lowers the number of free parameters in the inversion, resulting in a χ^2 hypersurface that should produce fewer local minima. In the coupled method, we invert the same number of data points as in the pixel-by-pixel method with fewer parameters. This adds more constraints to the inversion parameters and aids the inversion algorithm in finding the global minimum.

In the coupled method, a single value of the Marquardt parameter controls the step sizes for all parameters, whereas in the pixel-by-pixel method, each pixel has its own Marquardt parameter. This difference has consequences in effectively locating the global χ^2 minimum. A similar set of equations is solved in the PSF coupled inversion algorithm of [van Noort \(2012\)](#). In that paper, the author argues that to achieve convergence to the global minimum effectively, the inversion should be run for a small number of iterations (~ 10), after which the parameters are perturbed and used as initial values for the following inversion run. This way, the LM algorithm is kicked from any local minimum it may have strayed into to locate the global minimum.

Appendix C: globin

Current spectropolarimetric inversion codes like SIR ([Ruiz Cobo & del Toro Iniesta 1992](#)), SPINOR ([Solanki 1987](#); [Frutiger et al. 2000](#)) and NICOLE ([Socas-Navarro et al. 2015](#)) can invert both atmospheric and atomic parameters such as $\log(gf)$, the central wavelength of line, elemental abundance and line broadening parameters. However, these parameters are allowed to vary freely at every pixel in these codes. Here, we will describe in detail the implemented functionality of the `globin` code in which we have implemented the new method for retrieving atomic parameters.

Depth variations of physical parameters in an atmospheric model are commonly assumed to be a function of the optical depth, calculated at a reference wavelength of 5000 Å. To properly synthesize different spectral lines (e.g. NUV lines), we require the stratification of atmospheric models from photospheric up to chromospheric layers to be specified on a fine grid. Inverting observed spectra by iterative adjustment of all atmospheric parameters at each depth is very complicated due to the large number of free parameters. A feasible inversion of atmospheric structure, therefore, requires an approximation of the stratification of atmospheric parameters.

One kind of approximation is to parameterize the atmospheric stratification using a few points at properly chosen depths and interpolate between them to obtain values on a finer grid as needed for the spectral synthesis ([Ruiz Cobo & del Toro Iniesta 1992](#); [del Toro Iniesta 2003](#)). These points are called nodes and are placed over the range over which the line and its nearby continuum respond to changes in the atmosphere. In `globin`, nodes contain values of parameterized atmospheric parameters. Node positions need to be specified on an optical depth scale and can be chosen for each parameter independently. The interpolation is done either using non-overshooting Bezier interpolation polynomials of 2nd or 3rd order ([de la Cruz Rodríguez & Piskunov 2013](#)) or cubic splines under tension. The polynomial approximation, interpolation degree, and spline tension are chosen by the user. The values of the parameters from the highest node to the top boundary of the atmosphere are linearly extrapolated. The same procedure is also applied from the deepest node to the bottom boundary. For the temperature, we linearly extrapolate using the temperature gradient from the FAL-C model ([Fontenla et al. 1993](#)) for the lowest node ([Milić & van Noort 2018](#)). The temperature extrapolation to the

Table C.1. The parameter perturbations δp used for computing numerical response functions.

parameter	perturbation
T	1 K
v_{LOS}	1 m/s
v_{mic}	1 m/s
B	1 G
γ	0.01 rad
ϕ	0.01 rad
$\log(gf)$	0.001
$\Delta\lambda$	1 mÅ

higher layers has a lower limit of 2800 K in the highest point in the atmosphere. The same limit is also imposed on all temperature nodes. For lower temperatures, most hydrogen atoms form H^2 molecule, significantly reducing the population of neutral hydrogen atoms. This affects the continuum opacity contributions, such as from H^- ion, which produces a significantly different continuum level than the observed one.

The most time-consuming part of any inversion is the computation of the response functions. In the LTE approximation, the response functions can be computed analytically at the same time when the radiative transfer equation is solved (Ruiz Cobo & del Toro Iniesta 1992). In `globin`, we opted for a straightforward brute-force method and compute response functions using a numerical central difference scheme (see e.g. Quintero Noda et al. 2016). From the initial value of the inversion parameter, we make a positive perturbation by a small amount δp_k (see Table C.1), interpolate to obtain a refined stratification and compute the resulting Stokes spectrum \mathbf{S}^+ . Next, starting from the same initial value, we repeat the process with a negative perturbation of the same magnitude and compute the corresponding Stokes spectrum \mathbf{S}^- . The response function of a given parameter is computed as $\frac{\mathbf{S}^+ - \mathbf{S}^-}{2\delta p_k}$. This numerical approach to computing the response function allows the method to be applied arbitrarily to complex atmospheric stratifications.

The perturbation values given in Table C.1 are chosen to be small enough to approximate the response functions as a first-order perturbation of the spectrum but large enough to produce differences in the spectra that are significantly larger than any numerical uncertainty. In the case of much larger perturbations of the physical parameters, we would reach a non-linear perturbation of the spectrum due to the non-linearity of the radiative transfer equation.

The values of the response functions for different physical parameters span several orders of magnitude. The origin of this lies in the sensitivity of the Stokes spectra to perturbations of particular physical parameters and the choice of units for these parameters. It leads to unequal scaling between parameters, which influences the convergence properties of the LM algorithm. To improve the convergence of the inversion, Marquardt (1963) suggested scaling the response function of each k -th parameter with:

$$s_k = \sqrt{\sum_{i=1}^N J_{i,k}^2}, \quad (\text{C.1})$$

where the index i goes over all wavelengths for all four Stokes components. Dividing the computed response function with this scaling parameter yields a dimensionless response function. The Hessian matrix then becomes a matrix of parameter correlation coefficients (diagonal elements are equal to $1 + \lambda_M$). The parameter correction Δp_k can be converted back to proper units by multiplying it with the s_k .

During each iteration of the inversion process, the atmospheric model is assumed to be in hydrostatic equilibrium, so the temperature alone is sufficient to derive gas and electron pressures assuming the ideal gas law. From these, we obtain the hydrogen populations in the LTE approximation, necessary to compute spectra using the RH code.