ISSN 1845–8319

PARALLELIZATION OF THE SIR CODE

S. THONHOFER[1,2], L.R. BELLOT RUBIO[2], D. UTZ[1,2], J. JURČÁK[3], A. HANSLMEIER[1], I. PIANTSCHITSCH[1], J. PAURITSCH[1], B. LEMMERER[1] and S. GUTTENBRUNNER[1]

[1]*IGAM/Institute of Physics, University of Graz, Universitätsplatz 5, 8010 Graz, Austria*
[2]*Instituto de Astrofísica de Andalucía (CSIC), Apdo. de Correos 3004, 18080 Granada, Spain*
[3]*Astronomical Institute of the Academy of Sciences of the Czech Republic, Fričova 298, 251 65, Ondřejov, Czech Republic*

Abstract. A high-resolution 3-dimensional model of the photospheric magnetic field is essential for the investigation of small-scale solar magnetic phenomena. The SIR code is an advanced Stokes-inversion code that deduces physical quantities, e.g. magnetic field vector, temperature, and LOS velocity, from spectropolarimetric data. We extended this code by the capability of directly using large data sets and inverting the pixels in parallel. Due to this parallelization it is now feasible to apply the code directly on extensive data sets. Besides, we included the possibility to use different initial model atmospheres for the inversion, which enhances the quality of the results.

Key words: photosphere - spectropolarimetry - magnetic field

1. Introduction

For the investigation of small-scale magnetic phenomena in the solar photosphere, e.g. magnetic bright points (see Utz *et al.* 2009, 2010), knowledge about the underlying physical conditions is required. Our main interest focuses on determining the magnetic field vector and the stratification of the line-of-sight (LOS) velocity. The state-of-the-art method for deriving information about magnetic field from observational data is spectropolarimetry (for a comprehensive review see the book by del Toro Iniesta, 2003).

By means of spectropolarimeters (slit based spectrographs or magnetographs), the Zeeman-induced polarization of spectral lines is measured. In the next step, the so-called data inversion, the physical quantities are deduced from the Stokes spectra by solving the radiative transfer equation (RTE) numerically. Several computer programs for this mathematical inversion problem using different assumptions exist. Their common approach is to synthesize Stokes spectra from an atmospheric starting model and vary it until the best fit between the obtained synthetic spectra and the observed ones is reached.

We chose to use the SIR code (Stokes Inversion based on Response functions, Ruiz Cobo & del Toro Iniesta 1992) as it considers an atmospheric stratification with height, resulting in a 3-dimensional model of the photosphere. It assumes local thermodynamic equilibrium (LTE), which facilitates the inversion. Since its introduction, the SIR code has proven its flexibility (with respect to different spectral lines, different observational data, different atmospheric models) and its robustness, which resulted in a widespread usage among the community and a large amount of publications referring to it (NASA ADS lists 362 references to Ruiz Cobo & del Toro Iniesta 1992).

The SIR code in its basic configuration can invert only one pixel at once, complicating its usage on current spectropolarimetric data. This is especially true for inexperienced users. High-resolution data from modern instruments (such as Hinode/SOT, Sunrise/IMaX, SST/CRISP, GREGOR) usually provide a large amount of data consisting of thousands or millions of pixels (due to the high spatial resolution), especially when using time series. Therefore we decided to extend the SIR code in such a way that it is able to handle a complete image of spectropolarimetric data at once making it also necessary to improve the runtime performance. This leads to a mandatory upgrade of the basic code into a fully parallelized one as described in the following.

2. Methods

The original SIR code, which is written in FORTRAN, reads its input data (e.g. observed Stokes spectra, atomic parameters and abundances, initial guess model) from ASCII-files, performs the inversion, and writes the results (synthetic Stokes spectra of the best fit, atmospheric model and its errors, one each per inversion cycle) again to ASCII-files. The inversion is performed using a Marquardt algorithm (Marquardt 1963) for the fitting process. In order to decrease the computational effort, the physical quantities are only computed on discrete nodes in the vertical direction. The number of nodes for each parameter corresponds to the degree of freedom of the fitting problem and has to be specified by the user, thus it is one of the most important user-specified settings of SIR inversions. The usual approach is to perform 3 inversion cycles with an increasing number of nodes. For more details about the workflow and the numerics see Ruiz Cobo & del Toro Iniesta (1992) and Bellot Rubio (2003).

In order to be able to read and write the typically large amount of data (e.g. for Hinode SOT/SP 1024×37 pixels per scan, each with 112 spectral points, for the 4 Stokes parameters) we decided to use the standardized and popular FITS (Flexible Image Transport System) format for the upgraded version of the SIR code. We included the FITS input/output capability to our code using the CFITSIO library (Pence 1999).

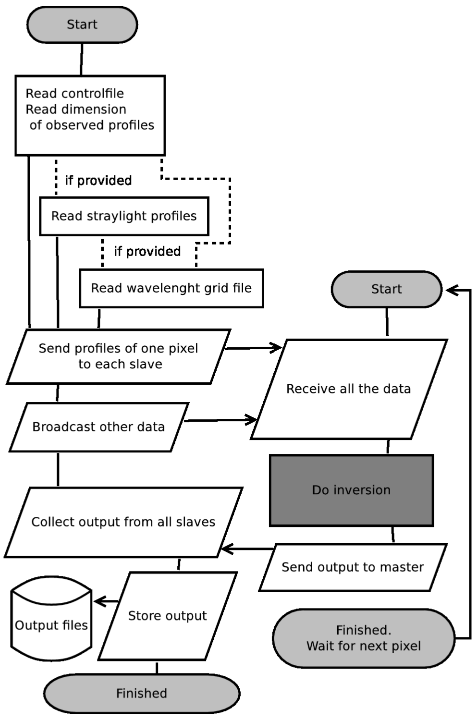

Figure 1: Workflow of the parallelized SIR code. The tree on the left side represents the master process, whereas the tree on the right stands for one of the slave processes. The box labelled “Do inversion” represents the original, serial, SIR code, detracted from all input/output routines and converted to a subroutine

The SIR code is designed to invert only one pixel per call, therefore it is possible to parallelize it, i.e. to invert several pixels in parallel on different computing nodes (e.g. CPU cores). Due to its complexity (the code consists of approximately 100 source files) and its stability we did not want to change the existing code comprehensively. Therefore we used MPI (Message Passing Interface, see Message Passing Interface Forum 1998), a commonly used technique for distributed memory architecture, for the parallelization. MPI defines a standard for the communication between the computing nodes. In our case there is no communication needed between two inverting nodes because each pixel is inverted separately, so we used a master-slave topology. This means that one master node performs all input operations, distributes the data to the slave nodes, collects the results from them, and stores the output. The slave nodes on the other side only need to call the existing SIR code, which we converted for reasons of lucidity to a subroutine and eliminated all input/output from. The workflow of the parallelized SIR code is schematically shown in Figure 1

In order to enhance the quality of the inversion results we furthermore included the possible usage of several different initial model atmospheres. If the user provides a file containing a set of different initial values and their gradients, the program creates atmospheric models containing all permutations of these values. Synthetic profiles are then calculated for each pixel using these initial models and the result with the best fit to the observed profiles (i.e., the lowest χ^2) is stored.

3. Results

Figure 2 shows the speedup of the parallelized SIR code compared to the original, serial version. The speedup is defined as $\frac{t_1}{t_n}$, where t_1 is the execution time of the program using one single CPU core, while t_n specifies the execution time using n cores. For the current study t_n was obtained on the IBM iDataPlex dx360 M3 cluster (38 nodes in total, each with 2x Intel Westmere Xeon E5645 6 Core, 2.40 GHz, 12 MB Cache) of the University of Graz. For t_1 we used in this case the execution time of the original SIR code for one pixel, averaged over 100 pixels, executed on the same cluster but using only one core. The ideal speedup should increase linearly with the number of processors used.

The usage of different initial model atmospheres for each pixel unfortunately increases the computation time significantly. Compared to the inver-

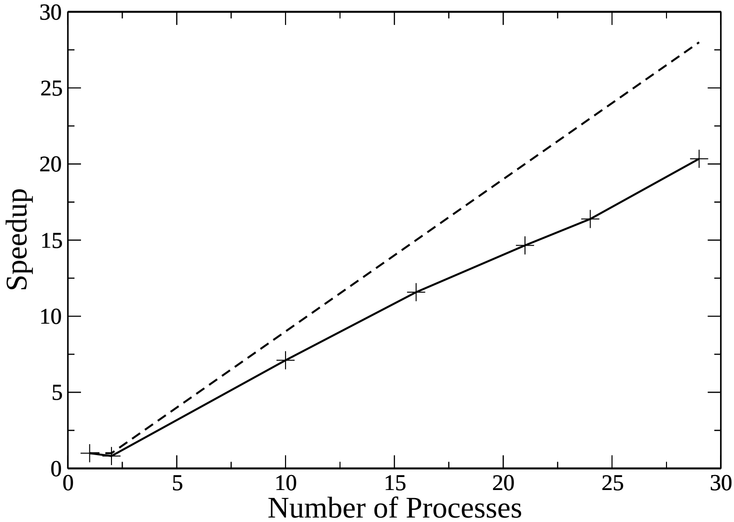

Figure 2: Speedup of the parallelized SIR code. The reference value for 1 process is the serial SIR code. Due to the Master-Slave topology the number of inverting processes is the number of processes minus 1. The dashed line marks the theoretical ideal speedup.

sion with only one initialization, the usage of 2 different values for e.g. magnetic field strength, its gradient, the gradient of the LOS-velocity and the gradient in inclination would last approximately 16 times longer ($2^4 = 16$). By systematically covering the whole range of reasonable initial models we can make sure that we find an inversion with a best fit solution. Experience showed that the SIR code converges well to a satisfying solution if the polarization signal is strong, i.e. the magnetic field strength is high. First test runs with quiet-sun data (Hinode, SST) already showed that the mean values of χ^2 over all pixels is lower when using at least two different models and one inversion cycle than using one initialization and performing three cycles.

4. Discussion

We extended the already existing SIR code by the capability of inverting extensive data sets of current high-resolution instruments more efficiently. This is done by changing the input and output format to the easy-to-use FITS format. Additionally, a commonly used parallelization technology increases the performance of this new version of the SIR code and enables it to be run on cluster systems. In practice, the speedup is limited by non-parallelizable parts of the program (e.g. input/output, communication between nodes). This fact is summarized in Amdahl's law (Amdahl 1967). We assume that the parallelized SIR code scales about one third worse than the ideal case due to the additional communication time. Further in-depth analysis might lead to the discovery of other limiting operations.

We included the possible usage of a new inversion scheme by subsequently repeating the inversion of a pixel using different initializations. If the user of our program performs the inversion in this way, better results, especially for regions of low magnetic field strengths, can be expected. It is essential that this technique is investigated and tested in more detail, e.g. by a detailed case study.

Acknowledgements

This work was funded by the Austrian Science Fund (FWF): P23618 (Dynamics of Magnetic Bright Points), the LLP ERASMUS program of the E.C. and the Förderungsstipendium of the University of Graz. D.U. wants to emphasize the special support given by project J3176 (Spectroscopic and Statistical Investigations on MBPs). Hinode is a Japanese mission developed and launched by ISAS/JAXA, with NAOJ as domestic partner and NASA and STFC (UK) as international partners. It is operated by these agencies in co-operation with ESA and NSC (Norway). J.J. is thankful to the ÖAD (Österreichischer Austauschdienst) and the MŠMT (Ministry of Education, Youth and Sports, Czech Republic) in supporting research stays in Austria. S.T., D.U., I.P., S.G. and B.L. are also thankful to the ÖAD and the MŠMT in supporting research stays in the Czech Republic. The authors gratefully acknowledge support from NAWI Graz.

References

Amdahl, G.: 1967, *AFIPS Press* **30**, 483.

Bellot Rubio, L.R.: 2003, Inversion of Stokes Profiles with SIR, (Freiburg: Kiepenheuer-Institut für Sonnenphysik)

del Toro Iniesta, J.C.: 2003, Introduction to Spectropolarimetry, Cambridge University Press, Cambridge, UK.

Marquardt, D.W.: 1963, *Journal of the Society for Industrial and Applied Mathematics* **11**, 431.

Message Passing Interface Forum, 1997, MPI2: Extensions to the Message Passing Interface.

Pence, W., 1999, *ASP Conf. Ser.* **172**, 487.

Ruiz Cobo, B. and del Toro Iniesta, J.C.: 1992, *Astrophys. J.* **398**, 375.

Utz, D., Hanslmeier, A., Möstl, C., Muller, R., Veronig, A., and Muthsam, H.: 2009., *Astron. Astrophys.* **498**, 289.

Utz, D., Hanslmeier, A., Muller, R., Veronig, A., Rybák, J., and Muthsam, H.: 2010., *Astron. Astrophys.* **511**, 39.