

IDEN2—A program for visual identification of spectral lines and energy levels in optical spectra of atoms and simple molecules[☆]

V.I. Azarov^a, A. Kramida^{b,*}, M.Ya. Vokhmentsev^c

^a Department of Research and Development, Parametric Technology Corporation, 140 Kendrick Street, Needham, MA 02494, USA

^b National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

^c ITM LLC, 59 Proletarskaya street, 640000 Kurgan, Russia

ARTICLE INFO

Article history:

Received 23 May 2017

Received in revised form 8 November 2017

Accepted 15 December 2017

Available online 24 December 2017

Keywords:

Atomic and molecular spectra

Energy levels

Line identification

Visual tool

ABSTRACT

The article describes a Java program that can be used in a user-friendly way to visually identify spectral lines observed in complex spectra with theoretically predicted transitions between atomic or molecular energy levels. The program arranges various information about spectral lines and energy levels in such a way that line identification and determination of positions of experimentally observed energy levels become much easier tasks that can be solved fast and efficiently.

Program summary

Program title: IDEN2

Program Files doi: <http://dx.doi.org/10.17632/mwr2tk84sw.1>

Licensing provisions: GPLv3

Programming language: Java v.1.8

Nature of problem: Line identification problem in atomic and molecular spectra that have a limited number of previously known low-excited energy levels and a large number of unidentified spectral lines (emission or absorption). The line identification consists of finding the lower and upper energy levels of transitions corresponding to each observed spectral line.

Solution method: For each energy level (experimentally known or unknown), the list of observed spectral lines is displayed on a wavenumber scale as an image. This image is repeated with certain vertical shifts in several columns corresponding to other levels connected with the investigated one by theoretically predicted transitions. In each such column, the image of the line spectrum is shifted by the value corresponding to the energy of the connecting level. Then the problem of location of an unknown level is reduced to scrolling the screen until several spectral lines in different columns align along the same horizontal line. Various gadgets provide complete information about all pertinent entities, such as the level identities, wavelengths, observed and calculated intensities.

Additional comments including restrictions and unusual features: A practical limitation on the size of the problem is imposed by the available RAM, which is typically sufficient to treat spectra with up to 100 000 spectral lines and 10 000 energy levels. The visual appearance of all gadgets, including font sizes, colors, and window sizes and positions, is completely customizable. An extensive Help system and User Guide are included.

Published by Elsevier B.V.

1. Introduction

The problem of identification of spectral lines is ubiquitous in atomic spectroscopy. If no energy levels are known in a

spectrum, the identification usually starts with a search for repeating wavenumber differences in pairs of observed lines. Such line pairs are prime suspects to originate from one upper energy level and terminate on two lower levels common to all these line pairs. Alternatively, one can look for the strongest predicted transitions in the spectrum and try matching them with the strongest observed lines. Once a few energy levels have been established, the search continues using a more general Ritz combination principle for transitions forming closed loops in the graph of transitions between predicted and observed levels. However, even if a Ritz combination of observed wavenumbers of several lines is satisfied

[☆] This paper and its associated computer program are available via the Computer Physics Communication homepage on ScienceDirect (<http://www.sciencedirect.com/science/journal/00104655>).

* Corresponding author.

E-mail addresses: vazarov@ptc.com (V.I. Azarov), alexander.kramida@nist.gov (A. Kramida).

almost perfectly, it may be due to a random coincidence and does not guarantee correct identification. Thus, a more sophisticated analysis is required to confirm the identification. In this analysis, the observed wavenumbers, line intensities, and positions of tentatively identified energy levels are compared with theoretical predictions. In complex spectra with hundreds of energy levels and thousands of possible transitions between them, this task is extremely difficult and time-consuming.

In the early 1990s, one of the present authors (VA) suggested a new formal approach to the line-and-level identification problem [1] and implemented it in a Fortran code IDEN [2] at the Institute of Spectroscopy, Troitsk, Russia.

By 1999, the code had been used in Troitsk (Russia), St. Francis Xavier University (Canada), Paris-Meudon Observatory (France) and Universiteit van Amsterdam (The Netherlands) spectroscopic laboratories in many analyses of complex spectra of atomic ions. These analyses included the spectra that were the most challenging in the mid-1990s: Os V and Os IV [3,4]. Since then, many 5d spectra have been studied utilizing the code. Some examples are the analyses of the Ta III, Ir IV, Pt IV and Pt V [5–8] spectra, each having a few hundred established levels and up to a few thousands identified spectral lines. Success in the analyses of these spectra was largely due to strength of the approach implemented in the program IDEN. Although the original version of the code had a big impact on the progress in studying 5d elements, it was restricted to Windows-based computers, and it had limitations on the number of included energy levels (<1000), transitions (<5000) and spectral lines (<10 000).

In 1999, the second of the present authors rewrote the code using the Delphi¹ programming system (again for Windows-based computers only). This version of the code was called IDEN1. In 2006, it was extended to allow analyses of spectra of simple diatomic molecules (see, e.g., Roudjane et al. [9]). It could handle up to 10 000 energy levels and 50 000 transitions and was used in several analyses of atomic spectra at the National Institute of Standards and Technology (NIST), Gaithersburg, USA (e.g., Kramida et al. [10]). However, all these code versions were unpublished, limiting their applications.

The purpose of the present work was to rewrite the code in the Java programming language, making it readily available and customizable on any computer system supporting Java.

2. The basics of the method

The basic idea of this method is to search for new energy levels, one level at a time, by surveying the Ritz combinations involving all known levels connected with the given one by a predicted transition. To facilitate this search, both the experimental data (energy levels, wavenumbers and intensities of observed spectral lines) and theoretical predictions (approximate positions of predicted levels and transition intensities) are combined and organized on a computer screen in a compact and convenient way. All transitions originating from the selected level are displayed as vertical bands sorted from left to right in the order of decreasing predicted intensity. Each such band contains a copy of the entire range of observed wavenumbers displayed as colored rectangles. The height of each rectangle corresponds to the measurement uncertainty of the line. The dispersion (from top to bottom) is the same in each band, but each such copy of the spectrum is shifted vertically. The shift is defined by the energy of the connecting level, different for each transition. The color of the displayed spectral-line rectangles reflects important properties of the lines. In particular, unidentified

lines have a green color, lines already identified as transitions between some known levels are painted in magenta, and lines identified as transitions from the selected level to the connecting level of the band are painted red.

Then the task of identifying an unknown level reduces to scrolling the entire screen vertically, which changes the assumed energy of the selected level. The scrolling continues until such a position is found where line-image rectangles in several transition bands align near a horizontal line at the center of the screen. This position satisfies the Ritz combination principle for transitions to the corresponding levels. The lines participating in these Ritz combinations should normally be located at the left part of the screen, corresponding to the strongest predicted transitions. Information about the observed and predicted intensities, wavelengths, and other properties of all displayed energy levels, lines, and transitions is readily available on the screen and needs to be checked in order to verify that the found lines can indeed be assigned to transitions in their bands. Once verified, the level position is fixed at the newly found energy value, and the identified lines are painted red. All these operations are made with the computer mouse. Thus, the user's task is reduced to scrolling the screen (occasionally zooming in and out) and painting as many green lines as possible in red color.

3. Differences from other methods and codes

There are no other published codes implementing the method used in IDEN2. However, other approaches to the task of identification of unknown energy levels and transitions in atomic spectra do exist. As noted in the Introduction, the most common method is to search for repeating differences in wavenumbers of observed spectral lines. Such search is often assisted by computer codes. However, none of such codes offer visual organization of all relevant data in such an efficient way as implemented in IDEN2. With that old method of searching for repeating wavenumber differences, data from theoretical calculations, which are essential tools in the search for new identifications, could not be combined with the list of observed wavenumbers by any means other than writing notes on pieces of paper and using a calculator. In contrast, in IDEN2 theoretical information is efficiently combined with experimental data in layouts of graphical screens, each energy level and each transition can be commented within the code, and these comments are easily accessible.

Another commonly used approach to the line-and-level identification problem employs the so-called square arrays. This method requires that the entire set of predicted energy levels is divided in smaller subsets, and transitions are sought between levels belonging to two such subsets. Thus, transition arrays are investigated for each pair of level subsets separately. Such transition arrays are depicted as square tables with upper levels given in the columns and lower levels in the rows of the table. In each cell of the table, one would have to write theoretically predicted intensity, wavenumber, and/or wavelength and record experimental data for observed lines considered as possible candidates for these transitions. This method involves a tedious work of checking each Ritz combination by hand, which is time-consuming and error-prone. In addition, complex spectra often involve energy levels that have transitions both from and to them, which make it impossible to separate the analysis of different transition arrays from each other. Moreover, it is not uncommon to have transition arrays involving many tens of energy levels. In such cases the square arrays become physically too large to handle. An attempt to computerize the work with square arrays was made by Reader [11]. However, unlike IDEN2, his set of codes does not automate the search for Ritz combinations and offers little assistance in keeping track of different possible level positions. In addition, large transition arrays do not fit to one screen in Reader's spreadsheets and are difficult to work with.

¹ Commercial products are identified in this paper for adequate specification of code versions. This identification does not imply recommendation or endorsement by NIST.

A specialized set of computer codes developed by Windholz and his team [12] includes a visualization code that uses a very different approach. However, it is restricted to the particular experimental methods used in that group, Fourier-transform and laser spectroscopy of hyperfine atomic spectra. Unlike IDEN2, it does not allow working with spectra that do not possess resolved hyperfine structure of the lines.

Compared with the previous versions of the IDEN code, IDEN2 has many enhancements, such as the ability to store and display comments to levels and transitions, navigable history of operations, an option to hide and show images for transitions that are irrelevant to the level search or obscure the view, options for sorting the line and level lists, editing the line list, and many other convenient features making it more efficient. Since none of the previous versions were published, we do not find it necessary to give a complete list of enhancements.

4. Details of the code operation

Details of working with the code are given in the extensive User Guide included with the package, which is also available as a navigable Help system accessible from the main code window. As an illustration, Fig. 1 shows a snapshot of the main window of the code. All the relevant concepts and terms, such as *energy level*, *transition*, *spectral line*, and *glued level* are explained in the User Guide.

The screenshot depicted in Fig. 1 is from the recently published analysis of the Cu II spectrum [13]. It shows the previously unknown $3d^8(^3P)4s4p(^3P^{\circ})^3S^{\circ}_1$ level (sequential number 1911) predicted by a theoretical calculation to be at $138\,992.8\text{ cm}^{-1}$ positioned by using mouse controls of the code at the best found experimental energy, $138\,516.51\text{ cm}^{-1}$. At this position, four green images of previously unidentified observed spectral lines corresponding to predicted transitions down to levels 2060, 2061, 2059, and 2062 are aligned at the horizontal axis (21) in the left part of the picture. Their observed intensities (numbers in rows of the picture footer 9) are decreasing from left to right, similar to the predicted intensities shown in the picture header (2). The green color of the line images represents the fact that these lines have not been previously assigned to any other transition. There is also a spectral line corresponding to a transition from the level 1911 to the level 2066 in column 7 of the picture aligned near the same horizontal axis. When the snapshot was taken, the mouse cursor was located within the rectangle of this spectral line. Thus, information about this line is shown at the bottom of the picture in the information block 13. The magenta color of the rectangle indicates that this particular line had already been assigned to another transition (between levels 1672 and 2059; see information block 13) with a much greater predicted intensity $I_{\text{calc}} = 141$ compared to the transition $2066 \rightarrow 1911$ (69; see the I_{calc} value in column 7 of the picture). Therefore, the fact that this line satisfies the Ritz combination principle should be treated as a spurious coincidence. Nevertheless, there are four observed spectral lines associated with transitions involving the level 1911, all having reasonable relative intensities and satisfying the Ritz combination principle. The latter fact is confirmed by the horizontal axis crossing the rectangular images of the spectral lines within their heights (corresponding to their measurement uncertainties). Scrolling the picture up and down confirms that there are no alternative positions of the level 1911 providing a similar alignment of other spectral lines. This verifies that this level is indeed located at $138\,516.510\text{ cm}^{-1}$. The identities of the connecting levels shown as numbers in the picture header can easily be revealed by invoking pop-up windows with a mouse click. Further steps in the work include “fixing” the level 1911 at the newly found position (see the “Fixed” checkbox in block 3) and “painting” the green line rectangles in the four

leftmost columns of the picture in red color by assigning the corresponding lines to transitions 1 through 4 (this is done by right-clicking the mouse at the rectangles of the line images and selecting the “Assign” command from the pop-up menu).

The level energies determined with the IDEN2 code should be treated as preliminary. They require additional level optimization, for which other computer codes exist (see references in [13]).

5. Pre-requisites

Besides having a computer with a good graphics display and sufficiently large RAM (≥ 2 Gbyte is recommended) and an operating system with Java, two important requirements must be satisfied in order to make a line-and-level identification problem solvable with IDEN2:

(1) A reasonably good theoretical prediction of the energy levels and transition rates must be available for the considered spectrum. These quantities are extensively used in the analysis. The exact meaning of “reasonably good” varies, depending on the spectrum. In some spectra, even a small modification of the theoretical model leads to drastic re-distribution of predicted transition intensities in the spectrum. Thus, an effort should be made to calculate the spectral data as accurately as possible.

(2) The list of observed lines should be as complete and accurate as possible. In atomic spectroscopy, the spectrum is often recorded in several pieces under different experimental conditions. The most accurate measurement should be selected for each observed line from those multiple spectrograms. Then the lines should be combined in a single list sorted by wavelength, making sure that each entry in the list corresponds to a unique spectral line. The measurement uncertainties should be evaluated for each line.

An effort should be made to reduce all observed line intensities to the same uniform scale. In the present version of IDEN2, the values of intensities must be integer numbers varying from 0 to 99 999. The dynamic range of photographic plates is usually about a factor of 100, while Fourier transform spectrograms, as well as phosphor image plates, may have a dynamic range of $\approx 100\,000$. Even with photographic plates, if there are multiple different exposures of the same spectral regions, the dynamic range of observed intensities can be as large as 10^6 . In such cases, it is reasonable to reduce the intensities to a logarithmic scale, so that their rounded values could be used in IDEN2. Although it is not required, it is better to remove possible variations of recording sensitivity with wavelength from the observed intensities, so that they could be quantitatively compared with theoretical predictions.

6. Input/output files

The program uses several input files, some of which are mandatory and some optional. The mandatory files are the list of observed lines *dlv.dat*, the list of predicted and observed energy levels *enlev.dat*, and the list of predicted transitions *trans.dat*. If wavelength uncertainties are not specified in the lines file *dlv.dat* for each line, an additional input file *numset.dat* is also required. This file defines the default wavelength uncertainties for specific wavelength ranges covering the entire range of lines in *dlv.dat*. Preparation of the line list is essentially manual (aided by any ASCII text editing software).

As noted in the previous section, both observed and calculated line intensities may cover a large dynamic range of many orders of magnitude. However, the IDEN2 code restricts the format of these values in the input files to integers from zero to 99 999. For visual clarity of data displayed in the program’s main window and gadgets, it is best to convert the line intensities to a logarithmic scale. Unknown observed intensities should be represented by zero values.

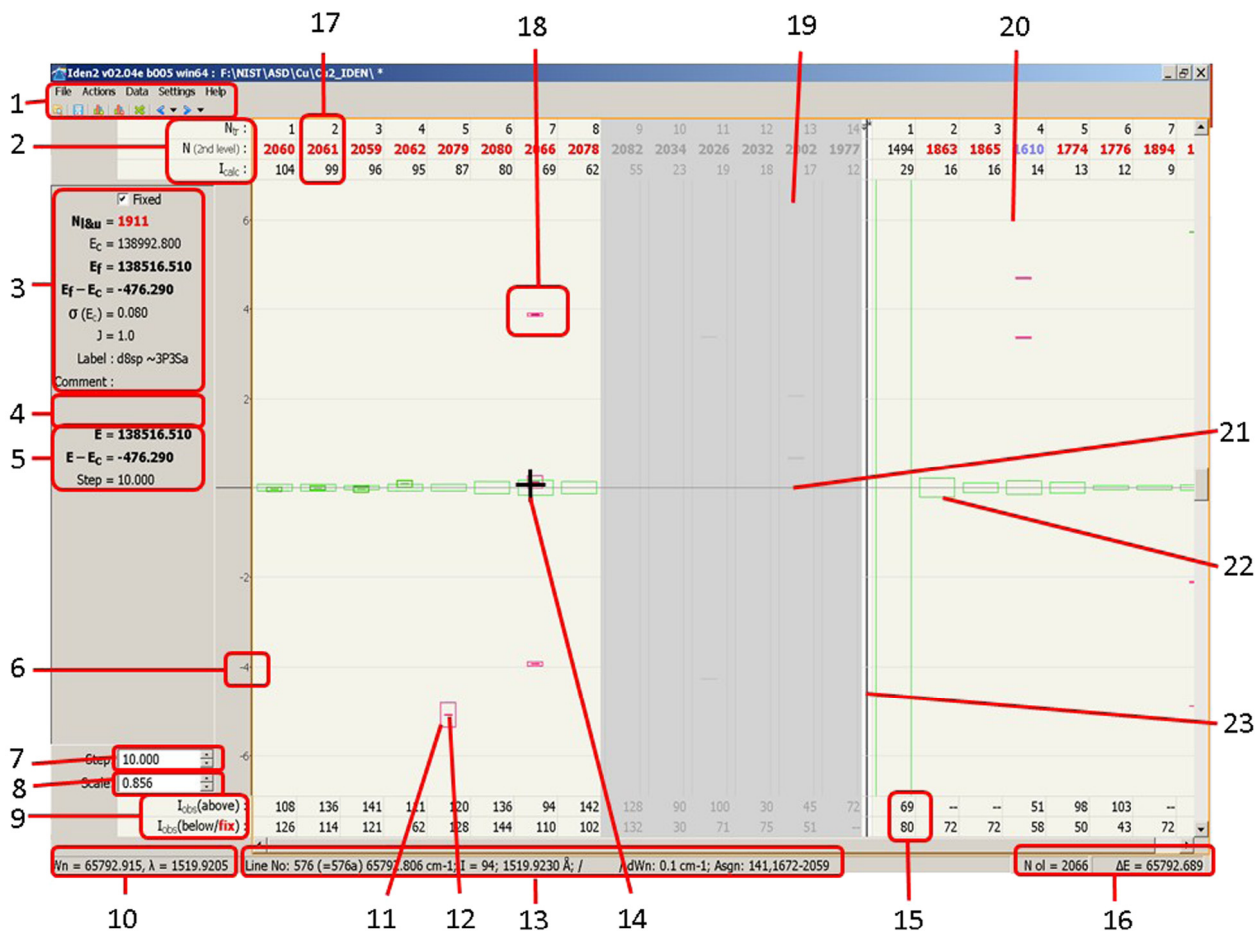


Fig. 1. Layout of the main window of the IDEN2 code. The labeled elements are explained below.

1. The main menu and the toolbar.
2. Labels designating the rows of the picture header.
3. Checkbox used to fix/unfix the energy level value and the information related to the selected level.
4. Label showing if the level is *glued* to another one (blank when the level is not glued).
5. Current position of the selected energy level, its difference from the calculated energy, and the current step value (in units of cm⁻¹).
6. A scale tick and its label.
7. Spinner controlling the step value.
8. Spinner controlling the picture scale.
9. Labels designating the rows of the picture footer.
10. Wavenumber and wavelength corresponding to the current position of the cursor.
11. A rectangle showing the uncertainty bounds of a line.
12. A dash in the middle of the line rectangle that shows the exact position of the wavenumber of the line.
13. Brief information about the spectral line under the mouse cursor.
14. The mouse cursor.
15. The footer cells show the intensities of the nearest lines above and below the x-axis (or the intensity of the line assigned to this transition, if there is one).
16. Sequential index number of the "other" (connecting) level of the transition and the transition wavenumber, i.e., the difference between the calculated or fixed energy of the current level and the "other" (connecting) level of the transition under the mouse cursor. If the level is fixed, its "fixed" energy is used in the calculation of the wavenumber; otherwise, the calculated energy is used.
17. The header cells show the sequential number of the transition shown in the column below, the index number of the "other" (connecting) level of the transition, and the calculated intensity of the transition.
18. A spectral line. The lines can be shown in different colors depending on the line properties.
19. Grayed-out "hidden" transition columns. Irrelevant transitions can be "hidden" for better visual clarity. They can be, e.g., too weak to be observable or have wavelengths outside of the observation range.
20. A scale line.
21. The x-axis line.
22. Uncertainty rectangle of the connecting level of a transition.
23. The divider between transitions from the selected level down to lower levels (on the left) and transitions from upper levels down to the selected one (on the right).

The lists of energy levels and transitions are usually computed using sophisticated theoretical atomic or molecular (AM) codes. Output files of those codes have various formats that are not standardized. However, it is not difficult to write custom utility codes converting the output of an AM structure code to the format of the input files required by IDEN2. Such a utility code, named *conv_out*, exists for the widely-used Cowan atomic structure code suite [14] and is included in Kramida's version of that suite [15]. Instructions

for usage of that suite of Cowan's codes, designed and compiled for the Windows operating system, are included in the package [15]. They also describe how to recompile it for any operating system other than Windows.

Details of the file names and formats, as well as the program concepts and operation instructions are explained in the User Guide included in the package, a copy of which is also accessible from within the program in the Help menu item.

7. Installing and running the code

No special actions are required to install the package, except installing the freely available Java runtime code if it is not already installed. The compiled code is in the executable file **Iden2.jar** in the `Iden2/Iden2_bin` directory of the distribution package. This file needs to be copied to the computer's file system. Under some operating systems, such as Linux or Mac OS, the file needs to be made executable (this property is set up by using the operating system tools). If the Java runtime environment is not already installed, it can be downloaded from <https://www.java.com> and installed following instructions therein.

Most operating systems have a graphical user interface having some kind of a file-management tool, e.g., Windows Explorer in MS Windows. The program can be executed from it by double-clicking on its file name. Alternatively, the program can be started from a command-line shell by executing the command

```
java -jar Iden2.jar
```

8. Program package contents

The program is packaged together with the source files, sample input data files, and documentation in a zip archive file. The archive contains one directory, **Iden2**, with two sub-directories, **Iden2_bin** (containing the executable file **Iden2.jar**) and **Iden2_Distr** (containing the source files and documentation). The directory **Iden2_Distr** contains also the file **readme.html** with detailed instructions for developers how to compile the source codes and explanations of the contents of the distribution package.

The sample input data files are in the subdirectory **Iden2_Distr/Data** of the distribution package. There are two sample data sets, one for an atomic spectrum (data from the study of

singly ionized copper [13]) and one for a molecular spectrum (data from the study of the D_2 molecule [16]).

Acknowledgments

V. Azarov was working under contract SB1341-12-CQ-0011 with the National Institute of Standards and Technology, Gaithersburg, MD 20899, USA. We gratefully acknowledge the kind permission of the authors of Refs. [13] and [16] to use their data in the distribution package and valuable discussions with Dr. W.-Ü L. Tchang-Brillet.

References

- [1] V.I. Azarov, *Phys. Scr.* 44 (1991) 528.
- [2] V.I. Azarov, *Phys. Scr.* 48 (1993) 656.
- [3] V.I. Azarov, A.J.J. Raassen, Y.N. Joshi, P.H.M. Uylings, A.N. Ryabtsev, *Phys. Scr.* 56 (1997) 325.
- [4] A.N. Ryabtsev, A.J.J. Raassen, W.-Ü.L. Tchang-Brillet, Y.N. Joshi, P.H.M. Uylings, V.I. Azarov, *Phys. Scr.* 57 (1998) 82.
- [5] V.I. Azarov, W.-Ü.L. Tchang-Brillet, J.-F. Wyart, F.G. Meijer, *Phys. Scr.* 67 (2003) 190.
- [6] V.I. Azarov, R.R. Gayasov, *At. Data Nucl. Data Tables* 108 (2016) 81.
- [7] V.I. Azarov, R.R. Gayasov, *At. Data Nucl. Data Tables* 108 (2016) 118.
- [8] V.I. Azarov, R.R. Gayasov, *At. Data Nucl. Data Tables* 108 (2016) 154.
- [9] M. Roudjane, F. Launay, W.-Ü.L. Tchang-Brillet, *J. Chem. Phys.* 125 (2006) 214305.
- [10] A. Kramida, C.M. Brown, U. Feldman, J. Reader, *Phys. Scr.* 85 (2012) 025303.
- [11] J. Reader, *Comput. Phys.* 11 (2) (1997) 190.
- [12] L. Windholz, *Phys. Scr.* 91 (2016) 114003.
- [13] A. Kramida, G. Nave, J. Reader, *Atoms* 5 (2017) 9.
- [14] R.D. Cowan, *The Theory of Atomic Structure and Spectra*, University of California Press/Berkeley, CA, 1981.
- [15] A. Kramida, PC Version of Cowan Codes for MS Windows, online at <http://das101.isan.troitsk.ru/COWAN>, 2012.
- [16] M. Roudjane, W.-Ü.L. Tchang-Brillet, F. Launay, *J. Chem. Phys.* 127 (2007) 054307.