Ridge regression energy levels calculation of neutral ytterbium ($Z = 70$)

Yushu Yu(余雨姝)[1,2], Chen Yang(杨晨)[1,2,†], and Gang Jiang(蒋刚)[1,2]

[1] *Institute of Atomic and Molecular Physics, Sichuan University, Chengdu* 610065, *China*
[2] *Key Laboratory of High Energy Density Physics and Technology, Ministry of Education, Chengdu* 610065, *China*

(Received 6 September 2022; revised manuscript received 29 November 2022; accepted manuscript online 16 December 2022)

In view of the difficulty in calculating the atomic structure parameters of high-Z elements, the Hartree–Fock with relativistic corrections (HFR) theory in combination with the ridge regression (RR) algorithm rather than the Cowan code's least squares fitting (LSF) method is proposed and applied. By analyzing the energy level structure parameters of the HFR theory and using the fitting experimental energy level extrapolation method, some excited state energy levels of the Yb I ($Z = 70$) atom including the 4f open shell are calculated. The advantages of the ridge regression algorithm are demonstrated by comparing it with Cowan code's LSF results. In addition, the results obtained by the new method are compared with the experimental results and other theoretical results to demonstrate the reliability and accuracy of our approach.

Keywords: atomic data, ytterbium, energy levels, ridge regression algorithm

PACS: 31.15.xr, 32.30.–r, 95.30.Ky **DOI:** 10.1088/1674-1056/acac08

1. Introduction

The atomic structure data (energy levels, transition probabilities, radiative lifetimes, hyperfine structure constants, the Landé g-factors, isotope shifts data) of high-Z atoms play fundamental and critical roles in atomic physics. Accurately measuring or calculating these characteristics has great significance for the understanding of atomic structure and their ionic states and the applications in some technical fields, such as biological imaging,[1] medical diagnosis,[2] atomic optical clocks,[3] *etc.*

In recent years, with the development of photoionization detection technology, substantial progress has been made in the spectral analysis of complex atomic systems.[4–11] However, from the National Institute of Standards and Technology (NIST) Atomic Spectra Database, there is no systematic experimental data for high-Z elements, especially lanthanides and actinides.[12] On the one hand, the high cost of sample preparation and high radioactivity bring much inconvenience in the experimental measurement. On the other hand, with the increase of atomic number, complex electron correlation effects and relativistic effects bring a high computational cost, which is the major obstacle to calculating the parameters of complex atomic systems theoretically in fundamental research.

When dealing with a large amount of experimental data, it is appropriate to use machine learning for data mining and modeling research instead of expressing physical equations explicitly in the traditional scientific paradigm. Therefore, data-driven methods are a good way to obtain more atomic structure data and improve the accuracy of high-Z atomic data, avoiding expensive *ab initio* calculations. Nowadays, widely used atomic structure packages can be divided into two categories. The first category includes programs that use non-relativistic with relativistic corrections methods. For example, the Cowan code[13] developed by Cowan in 1968 was rather primitive, which adopted the single-configuration radial wave function Hartree–Fock (HF) method. There are also ATSP2K[14] developed by Fischer *et al.* in 2007 using a non-relativistic multiconfiguration HF (MCHF) method, CIV3[15] developed by Hibbert, *etc.* The second category of *ab initio* packages employ the Dirac equation that considers relativistic effects into account, including the GRASP2K[16] *ab initio* package that adopted the multiconfiguration Dirac–Hartree–Fock method published by Grant *et al.* in 2013; the FAC[17] program with many-body perturbation theory developed by Gu; the MDFGME[18–20] program applys the multiconfiguration Dirac–Fock matrix element method by Desclaux's group, *etc.*

The packages using the multiconfiguration Hartree–Fock method, *e.g.*, ATSP2K, GRASP2K, *etc.* are preferable in calculating the low-lying states of light elements. However, for elements containing nd and nf electrons, these codes are the same or even less accurate than Cowan code. A major reason is the Cowan code performs semi-empirical approach with parametric fitting, the least squares method was adopted to fit the *ab initio* calculation results and the existing experimental data by an iterative procedure. As a result, many effects are incorporated into the *ab initio* calculation, such as core polarization effects, the relativistic and electron correlation effects. The semi-empirical method has been widely used because of its advantages, including a relatively simple calculation model

†Corresponding author. E-mail: yangchen@scu.edu.cn

© 2023 Chinese Physical Society and IOP Publishing Ltd

and absolute convergence of the iterative process, significantly improving the calculation accuracy. In 1996, Uylings and coworkers[21] proposed applying the orthogonal operator method in spectral analysis. The orthogonal operator method can be considered an extension and improvement of the traditional least squares fitting (LSF) method, introducing many small parameters to account for specific interactions. The orthogonal operator method has many advantages. It uses the nonrelativistic approximation with relativistic corrections as a perturbation method, the same as the Cowan code. In addition, small second-order interactions is also considered in the orthogonal operator method. However, the orthogonal operator method is only applicable to the d^n, $d^n l$, $d^n l^n l^2$, and f^2 configurations so far; for f^n configurations with $n > 2$, the method remains to be developed. Since 1990, the Cowan code has been cited approximately 150 times per year, covering 58 research categories ranging from optics, atomic physics, solid state physics, astronomy, chemistry to nuclear physics.[22] Among them, his contribution to the calculation of the spectra of transition elements, including lanthanides and actinides is particularly outstanding. Liggins *et al.*[23] used Cowan code calculated transition probabilities for parity-forbidden Mn II lines and the mean absolute error (MAE) between computed and experimental levels is 241 cm^{-1}. Saloman and his co-works reported neutral vanadium[24] and singly ionized vanadium[25] energy levels. Parametric calculations of the free ion spectrum of Er IV have been performed using the Cowan codes by Chikh *et al.*,[26] has led to the determination of 168 new energy levels in Er IV, in addition to the 120 previously known ones. Inspired by the Cowan code, if the current state-of-the-art data-driven machine learning (ML) method is integrated into the calculation of atomic structure parameters, the accuracy of the current atomic structure calculation program can be improved.

Ytterbium ($Z = 70$) is one of the typical lanthanide elements, and its neutral atom ground state is $4f^{14}6s^2\ {}^1S_0$. Two energy level systems are generated about the neutral Yb excited states: The $4f^{14}nln'l'$ configuration excited by 6s shell electrons and the $4f^{13}nln'l'n''l''$ configuration level system excited by $4f$ shell electrons. Yb I has been recognized by some groups as an excellent atomic optical clock candidate because of its narrow $4f^{14}6s^2\ {}^1S_0 \rightarrow 4f^{14}6s6p\ {}^3P_0$ transition.[27] Energy levels and Landé g-factors, as the most basic features of atomic structure spectroscopy, have been studied for a long time. King has measured the spectrum of neutral Yb atoms using an electric furnace in 1931.[28] Meggers and his colleagues studied atomic spectra of rare earth elements, reporting the wavelengths and estimated intensities of 1791 lines of neutral ytterbium in the region from 2155 Å to 31308 Å.[29] Recently, using microwave and radio frequency resonance methods, Niyaz *et al.* observed the transition of ytterbium from $6s(n+3)d\ {}^1D_2 \rightarrow 6snl$, $4 \leq l \leq 6$.[30] Besides experimental observations, many groups have developed new theoretical methods to calculate the structural parameters of Yb atoms. Ternovsky *et al.*[31] applied the method of relativistic many-body perturbation and the Dirac–Kohn–Sham zeroth approximation combined with the generalized relativistic energy method , and calculated the energy and width of the self-ionizing resonance of $4f^{14}7s6p$, $4f^{13}[{}^2F_{7/2}]6s^2np[5/2]_2$, $4f^{13}[{}^7F_{7/2}]6s^2nf[5/2]_2$. Dzuba and Flambaum[32] developed and used an efficient version of the multi-electron configuration interaction method, the fast configuration–interaction method (FCI), to calculate energy levels of some low-lying states of ytterbium.

In this work, we further develop a more efficient data-driven method that uses the ridge regression (RR) algorithm to replace the Cowan code's LSF method. Taking the relative energy levels of the Yb I atom as an example, we calculate the energy levels for the partially excited states of the 6s shell and 4f shell. The results from the LSF method and the RR algorithm are compared with the experimental values, illustrating the advantages of the RR model in extrapolative prediction. In addition, the HFR theory in combination with the RR algorithm was used to calculate the energy levels and Landé g-factors of some highly excited states of the Yb I atom, and reliable results were obtained.

2. Calculation methods

2.1. Ridge regression algorithm

Ridge regression[33] is an improved linear regression algorithm. Unlike the unbiased estimation of linear regression, ridge regression acts more like an orthogonal system by adding a small positive quantity, which is more practical and reliable.

For a given dataset $\mathcal{D}$, as shown in the following equation,

$$\mathcal{D} = \left\{\left(\boldsymbol{x}^{(1)}, y^{(1)}\right), \left(\boldsymbol{x}^{(2)}, y^{(2)}\right), \ldots, \left(\boldsymbol{x}^{(i)}, y^{(i)}\right), \ldots, \left(\boldsymbol{x}^{(m)}, y^{(m)}\right)\right\}, \tag{1}$$

where m represents the size of the input dataset; $\boldsymbol{x}^{(i)}$ represents a set of input data, $y^{(i)}$ is the true values corresponding to the sample $\boldsymbol{x}^{(i)}$, and then start building a linear regression model to make predictions

$$h_\theta(\boldsymbol{x}) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n, \tag{2}$$

where θ_i represents the model parameters of linear regression and x_i represents input data. For convenience, $x_0 = 1$ is introduced. For each set of input data $\boldsymbol{x}^{(i)}$, the ultimate goal is to make the predicted value $h_\theta\left(\boldsymbol{x}^{(i)}\right)$ as close as possible to the true value $y^{(i)}$. To solve this problem, we define the loss function $J(\theta)$.

For least squares linear regression, its loss function is

$$J(\theta) = \sum_{i=1}^{m} \left(h_\theta\left(\boldsymbol{x}^{(i)}\right) - y^{(i)}\right)^2. \tag{3}$$

Using L_2 regularization, the loss function should be modified as

$$J(\theta)=\sum_{i=1}^{m}\left(h_\theta\left(\boldsymbol{x}^{(i)}\right)-y^{(i)}\right)^2+\lambda\sum_{j=0}^{n}\theta_j^2, \tag{4}$$

where λ is the regularization coefficient, which is generally related to specific problems such as training data and application scenarios, it is necessary to find a better option through tuning parameters. Using the L_2 regularized ridge regression model, taking into account the coefficient θ_j in the prediction function $h_\theta(\boldsymbol{x}^{(i)})$, it serves as a penalty mechanism to prevent overfitting of the lines.

2.2. HFR methods

The HFR method combined with the RR algorithm optimizes the Slater radial integral parameters to minimize the average deviation between the calculated energy levels and the selected experimental values. Thus, the required calculated energy levels are obtained. For an atomic system with nuclear charge number Z_0 and electron number N, in atomic units, the Hamiltonian is

$$\boldsymbol{H}=-\sum_i\nabla_i^2-\sum_i\frac{2Z_0}{r_i}+\sum_{i>j}\sum_j\frac{2}{r_{ij}}+\sum_i\zeta_i(r_i)\boldsymbol{l}_i\boldsymbol{s}_i, \tag{5}$$

where $r_i=|\boldsymbol{r}_i|$ is the distance of the i-th electron from the nucleus, $r_{ij}=\left|\boldsymbol{r}_i-\boldsymbol{r}_j\right|$ is the distance between the i-th and j-th electrons, and the summation over $i>j$ is over all pairs of electrons. The first term of the Hamiltonian represents the kinetic energy term of the electron, the second term is the potential energy term of the Coulomb interaction between the nucleus and the electron, the third term is the representation of Coulomb energy term between the electron and the electron, and the final term is the spin–orbit interaction term introduced by the relativistic correction. The proportionality factor $\zeta_i(r_i)$ mostly use the expression

$$\zeta(r)=\frac{\hbar^2}{2m^2c^2}\frac{1}{r}\left(\frac{\mathrm{d}V}{\mathrm{d}r}\right)$$

for a central field potential $V(r)=-q_Ne/r$.

The fitting calculation method is based on the Slater–Condon theory, which expands the unknown atomic wave function ψ_n with a set of known orthonormalized basis functions φ_b (calculated using Hartree–Fock or several more approximate methods),

$$\psi_n=\sum_b y_b^n\varphi_b. \tag{6}$$

The atomic wave function ψ_n and the energy E^n of the atom is obtained by solving the matrix equation: $\boldsymbol{H}\boldsymbol{Y}^n=\boldsymbol{E}^n\boldsymbol{Y}^n$ for every stationary quantum state n of interest.

According to the Slater–Condon theory[34] used by the HFR method, for the single configuration with the electron occupation number w_i of the orbital $n_il_i,(n_1l_1)^{w_1}(n_2l_2)^{w_2}\cdots(n_ql_q)^{w_q}$, its Hamiltonian matrix element is

$$\begin{aligned}\boldsymbol{H}_{ab}=&\delta_{ab}E_{\mathrm{av}}+\sum_{j=1}^{q}\left[\sum_{k>0}\left(f_k(l_jl_j)\right)_{ab}F^k(l_jl_j)+\left(d_j\right)_{ab}\zeta_j\right]\\&+\sum_{i=1}^{q-1}\sum_{j=i+1}^{q}\left[\sum_{k>0}\left(f_k(l_il_j)\right)_{ab}F^k(l_il_j)+\sum_k\left(g_k(l_il_j)\right)_{ab}G^k(l_il_j)\right]\\&\times\left(\left|l_i-l_j\right|\le k\le l_i+l_j\right),\end{aligned} \tag{7}$$

where E_{av} is the configuration average energy, $F^k(l_jl_j)$ is the Slater direct integral between equivalent electrons, $F^k(l_il_j)$ is the Slater direct integral between non-equivalent electrons, $G^k(l_il_j)$ is the Slater exchange integral between non-equivalent electrons, ζ_j is the radial integral related to the spin–orbit interaction. The f_k, g_k, and d_j are the angular coefficients of the radial integrals F^k, G^k, and ζ_j, respectively. Electrostatic and spin–orbit interaction parameters F^k, G^k, and ζ_j are defined in terms of the radial portion of a single-electron wave function $R(r)$ as the following equations

$$\begin{aligned}F^k(l_il_j)=&\int_0^\infty\int_0^\infty\frac{2r_<^k}{r_>^{k+1}}\\&\times R_i(r_1)R_j(r_2)R_i(r_1)R_j(r_2)r_1^2r_2^2\mathrm{d}r_1\mathrm{d}r_2,\end{aligned} \tag{8}$$

$$\begin{aligned}G^k(l_il_j)=&\int_0^\infty\int_0^\infty\frac{2r_<^k}{r_>^{k+1}}\\&\times R_i(r_1)R_j(r_2)R_i(r_2)R_j(r_1)r_1^2r_2^2\mathrm{d}r_1\mathrm{d}r_2,\end{aligned} \tag{9}$$

$$\zeta_j=\frac{\alpha^2}{2}\int_0^\infty\frac{1}{r}\left(\frac{\mathrm{d}V}{\mathrm{d}r}\right)\left|R_j(r)\right|^2\mathrm{d}r. \tag{10}$$

In the HFR method, the single-electron radial wave function for each specified electron configurations is first calculated using Hartree–Fock or several more approximate methods. Next, E_{av} and the Slater radial integrals for each configuration are calculated. Then, an energy matrix is set up for each possible value of J, and each matrix is diagonalized to obtain eigenvalues (energy levels) and eigenvectors. Furthermore, relativistic corrections are limited to calculations of mass velocity and Darwin corrections. The effects not considered in *ab initio*'s calculations are put into the ridge regression algorithm by treating the integral term in Eq. (7) as variable parameters. Our problem finally is to minimize the loss function

$$R(x_l)=\sum_i\left(E^i-T^i\right)^2+\lambda\sum_l x_l^2, \tag{11}$$

where E^i and T^i are the calculated values of HFR method and experimental energy level, respectively, x_l^2 is the square of radial integral value. We denote the various energy parameters $\left(E_{\mathrm{av}},F^k,G^k,\zeta_j\right)$ by the symbols x_l, and λ is the regularization coefficient. The various energy parameters are generally considered empirically adjustable quantities determined by the ridge regression fitted to the observed energy level.

3. Results and discussion

In this work, we calculated the relative energy levels and the Landé g-factor values for some excited states of the Yb I atom and compared them with the experimental data provided by NIST. In the calculation, the energy level values are relative to the ground state energy level $[Xe]4f^{14}6s^2$, and sorted in order from low to high energy levels. Taking the configuration $4f^{13}6s^26p$ excited by the 4f subshell as an example, we analyzed the advantages of the ridge regression algorithm, and then applied this method to highly excited states.

Table 1. Energy levels (in unit cm^{-1}) from experiment, *ab initio* calculations, and fitting calculations for Yb I $4f^{13}6s^26p$ configuration. The data marked with asterisks (*) represent the energy levels optimized by the LSF method and the RR algorithm.

					Fit I		Fit II		Fit III	
Configuration	Term	J	Expt.	*Ab initio*	Cowan code	This work	Cowan code	This work	Cowan code	This work
$4f^{13}6s^26p$	(7/2, 1/2)	3	32065.3	32277.2	32065.3*	32122.1*	31654.0*	32065.3*	32069.0*	32069.0*
		4	32273.6	32371.3	32273.6*	32216.6*	33456.0*	32273.6*	32280.3*	32280.3*
R_1[a]				165.0	0.0	56.9	885.2	0.0	5.4	5.4
	(7/2, 3/2)	5	35178.8	34455.3	32969.5	34355.9	34772.2*	35165.5*	35134.8*	35134.8*
		2	35197.0	34500.5	35023.8	34406.9	34839.4*	35210.3*	35205.4*	35205.4*
		3	35807.5	34790.8	33314.0	34690.0	39778.9	35482.9	35789.2*	35789.2*
		4	36061.0	34967.0	33471.0	34867.9	40314.2	35747.4	36104.5*	36104.5*
R_2				899.9	2111.7	996.7	2922.1	225.9	32.5	32.5
	(5/2, 1/2)	3		43213.3	46988.0	43331.2	42537.0	42416.3	42343.4	42343.4
		2	42531.9	43321.1	49249.4	43443.7	42196.1*	42531.9*	42521.9*	42521.9*
R_3				789.2	6717.5	911.8	335.8	0.0	10.0	10.0
	(5/2, 3/2)	3		45833.1	48265.1	46007.1	49006.2	45962.9	46218.5	46218.5
		1	44834.6	45122.9	47625.7	45297.5	40650.9	45258.9	44877.6*	44877.6*
		4	45497.6	45434.4	47998.0	45610.6	43712.0	45728.3	45473.3*	45473.3*
		2	45931.9	45654.2	48044.8	45828.1	48632.9	45776.1	45905.1*	45905.1*
R_4				234.0	2483.7	281.6	3054.4	293.0	32.4	32.4

[a] R_n is the root-mean-square error (RMSE), $R = \sqrt{\sum_i^N \frac{(E^i - T^i)^2}{N}}$ for each term.

Table 2. Slater parameters (in unit cm^{-1}) from *ab initio* calculations and fitting calculations for Yb I $4f^{13}6s^26p$ configuration, where ζ_f and ζ_p are radial integrals of the spin–orbit interaction for f shell electrons and p shell electrons, respectively. The radial integral $F^2(fp)$ represents the contribution to the Hamiltonian matrix elements of the direct Coulomb interaction between non-equivalent f and p electrons. $G^2(fp)$ and $G^4(fp)$ are exchange Coulomb interactions between non-equivalent f and p electrons.

		Fit I		Fit II		Fit III	
Parameters	*Ab initio*	Cowan code	This work	Cowan code	This work	Cowan code	This work
E_{av}	38565.6	39467.2	38565.6	39659.1	38755.0	38796.9	38796.9
ζ_f	3111.8	4243.5	3190.1	2320.3	2947.7	2903.0	2903.0
ζ_p	1514.7	625.1	1552.8	1170.4	2141.1	2136.2	2136.2
F^2 (fp)	2396.6	2102.2	2396.6	24744.7	2396.6	4474.6	4474.6
G^2 (fp)	566.1	6214.3	580.3	6412.3	566.1	1046.8	1046.8
G^4 (fp)	500.8	1211.0	513.4	0.0	1715.6	1027.2	1027.2

First, we use Cowan code's *ab initio* method to calculate energy levels of the configuration $4f^{13}6s^26p$. The results of all possible terms are listed in Table 1. To evaluate the accuracy of the calculations, for each term we introduce the root mean squared errors (RMSE): $R = \sqrt{\sum_i^N \frac{(E^i - T^i)^2}{N}}$, where T^i is the experimental energy levels from NIST, E^i is the theoretical levels, and N is the number of energy levels. The smaller the root mean square errors, the higher the accuracy. Clearly, Cowan code's *ab initio* method is almost acceptable. However, it is not highly accurate to identify the energy levels of lanthanide elements compared to the experimental results, with the RMSE ranging from 165.0 cm^{-1} to 899.9 cm^{-1} (Table 1). Such large R is because the interactions between atoms cannot be fully considered in Cowan code's *ab initio* calculation based on the HFR theory. To compensate for these shortcomings of Cowan code's *ab initio* method, we perform experimental values fitting using the LSF method and RR algorithms. The radial integrals $(E_{av}, \zeta_f, \zeta_p, F^2(fp), G^2(fp), G^4(fp))$ are considered simply as adjustable parameters. Those parameter values are determined empirically to give the best possible fit between the calculated eigenvalues and the experimental energy levels. With two experimental energy levels of the state $(7/2, 1/2)_{3,4}$ as fitted energy levels, we calculate the energy levels of configuration $4f^{13}6s^26p$ through Cowan code's LSF and RR methods, respectively. As shown in fit I of Table 1, Cowan code's LSF method can accurately fit the $(7/2, 1/2)_{3,4}$ energy levels, with R_1 of 0. It is a pity that the other energy levels are highly

inaccurate. The corresponding of R_2, R_3, R_4 is 2111.7 cm^{-1}, 6717.5 cm^{-1}, and 2483.7 cm^{-1}, obviously greater than that of *ab initio*, with the value of 899.9 cm^{-1}, 789.2 cm^{-1}, and 234.0 cm^{-1}. It should be stressed that the regularization term of the RR model limits the Hamiltonian parameters from being overly modified, resulting in more reasonable energy levels. The R, apart from the term $(7/2, 1/2)$, is all smaller than that of LSF and in good agreement with the results of Cowan code's *ab initio*. As reflected by R, our calculated results suggest that two experimental values cannot effectively improve the accuracy of energy level calculation. Therefore, five experimental energy levels are further considered. The resulting energy levels are displayed in fit II of Table 1. Comparing the data in fit I of Table 1, the R shows that the increased number of experimental levels helps optimize parameters and improve the accuracy of RR methods. The values of R_1, R_2, R_3, R_4 indicate that the results obtained from our developed RR method are optimal. Such evidence confirms that the RR model has a better generalization ability when experimental data are absent. The fitted calculation energy levels using Cowan code's LSF are wildly inaccurate due to the overfitting caused by the least squares method. Compared to the fit I of Table 1, the R is almost unchanged. Considering the number of adjustable parameters (here is 6), we considered at most 10 experimental energy levels to minimize the error caused by overfitting. As expected, the results in fit III of Table 1 reveal that the overfitting errors caused by the least squares method have been minimized when the number of experimental energy levels is up to 10. In other words, the calculation accuracy has been saturated. Thus, the result of the Cowan code is the same as that of this work in fit III. Moreover, as shown in fit II and fit III of Table 1, the variation in R suggests that the RR algorithm can provide more accurate results than Cowan code's LSF method when the observed energy levels are absent.

Table 3. Energy levels (in unit cm^{-1}) from experiment, *ab initio* calculations, and fitting calculations for Yb I $4f^{14}6snl$ configurations. The data marked with asterisks (*) represent the energy levels optimized by the LSF method and the RR algorithm.

					Fit I		Fit II	
Configuration	Term	J	Expt.	*Ab initio*	Cowan code	This work	Cowan code	This work
$4f^{14}6s^2$	^{1}S	0	0.0	0.0	0.0*	0.0*	0.0	0.0
$4f^{14}6s6p$	^{3}P	0	17288.4	11235.3	17269.9*	17288.4*	17288.4*	17288.4*
		1	17992.0	11668.9	17930.3	18009.0	17908.8	17908.8
		2	19710.4	12626.1	19924.4	20421.9	19710.4*	19710.4*
	^{1}P	1	25068.2	26035.3	25132.0*	25068.5*	25068.2*	25068.2*
R_1[a]				5651.2	116.2	355.8	32.2	32.2
$4f^{14}5d6s$	^{3}D	1	24489.1	20789.3	24489.1*	24489.1*	24489.1*	24489.1*
		2	24751.9	21257.1	24780.3	24834.2	24751.9*	24751.9*
		3	25270.9	22128.8	27069.3	26865.7	25275.2	25275.2
	^{1}D	2	27677.6	27658.7	27677.7*	27678.5*	27677.7*	27677.7*
R_2				2990.7	899.3	798.4	2.2	2.2
$4f^{14}6s7p$	^{3}P	0	38090.7	33330.9	38090.7*	38090.7*	38067.6*	38067.6*
		1	38174.2	33421.3	38174.2*	38174.2*	38204.9*	38204.9*
		2	38552.0	33639.9	38350.2	38373.0	38544.6*	38544.6*
	^{1}P	1	40564.0	35109.1	43119.1	39849.0	40563.7*	40563.7*
R_3				4978.2	1281.5	368.5	19.6	19.6
$4f^{14}6s6d$	^{3}D	1	39808.7	35330.3	39808.7*	39808.7*	39807.6*	39807.6*
		2	39838.0	35389.7	39838.0*	39838.0*	39839.3*	39839.3*
		3	39966.1	35526.8	61110.0	39887.1	39966.2*	39966.2*
	^{1}D	2	40061.5	35893.0	61154.1	40571.3	40061.3*	40061.3*
R_4				4385.4	14932.9	257.9	0.9	0.9

[a] R_n is the root-mean-square error (RMSE), $R = \sqrt{\sum_i^N \frac{(E^i - T^i)^2}{N}}$ for each term.

For the fitting calculation of configuration $4f^{13}6s^26p$ using different methods, we also analyzed the radial integral parameters of its Hamiltonian matrix elements (Table 2). Six slater parameters (E_{av}, ζ_f, ζ_p, $F^2(fp)$, $G^2(fp)$, $G^4(fp)$) are considered simply as adjustable parameters. The increased number of experimental levels results in well-defined parameters and improved accuracy of predicted energy levels. From the results given by the Cowan code, we find that some parameter values are too large or equal to zero (fit I and fit II). Significantly, the physical meanings of the slater parameters are trimmed excessively. The overfitting phenomenon caused by the least square method is the main reason that leads to these unreasonable parameters. Compared with Cowan code's LSF method, the parameter values given by the RR algorithm are more reasonable and stable. Consequently, Slater parameters with rational physical meaning can provide more accurate energy levels.

In Table 3, we also calculate the energy levels of $4f^{14}6snl$ configurations using different methods. We see that Cowan code's *ab initio* method results are not accurate enough. From

fit I, the results of RMSE can fully demonstrate the advantages of the ridge regression model. The RMSE of the RR algorithm is smaller than that of the LSF method when fitting the same number of experimental data. Importantly, more physically meaningful results have been obtained. For configuration $4f^{14}6s6d$, in fit I, we fit two experimental energy levels of $^3D_1$,$^3D_2$. The energy level results of terms $^1D_2$,$^3D_3$ from Cowan code's LSF method are 61154.1 cm^{-1} and 61110.0 cm^{-1}, respectively, which have very large deviations (about 2000 cm^{-1}) from experimental and *ab initio* results. The inaccurate results are caused by the overfitting. To keep the function prediction more consistent with the results of the added experimental data, large errors in the prediction results. The RR algorithm effectively avoids this overfitting, with R_4 is 257.9 cm^{-1}. Undoubtedly, the accuracy can also be greatly improved when more experimental levels are fitted, with the values of R_1, R_2, R_3, R_4 are less than 40 cm^{-1} (fit II).

Based on the ridge regression model, in the case of few data, the *ab initio* results can be modified to a certain extent for different excited states to optimize the Hamiltonian. We use this method to perform fitting calculations for energy levels and Landé g-factors of other excited states, and compared with other work, which are summarized in Tables 4 and 5 (odd and even parities, respectively). All the experimental energy levels taken from NIST of these excited states were included in the fitting procedure. As shown in Tables 4 and 5, our calculated levels fit well with the experimental results coming from NIST, with the absolute errors (ΔE), apart from individual levels (namely, $(7/2,3/2)_{2,3}$ and $(7/2,5/2)_{2,4,6}$ of $4f^{13}5d6s^2$), are less than 50 cm^{-1}. Interestingly, it can be seen that our results relative to the existing experimental data are distinctly better than the previous theoretical reports (Tables 4 and 5).

Table 4. Yb I odd-parity excited state fitting calculated energy levels (in unit cm^{-1}), Landé g-factor values, where ΔE (in unit cm^{-1}) and Δg are absolute errors between experimental and theoretical energies and Landé g-factors, respectively.

			Level				Landé g		
Configuration	Term	J	Expt.[12]	This work	Other work	ΔE	Expt.[12]	This work	Δg
$4f^{14}6s6p$	^{3}P	0	17288.4	17288.4	17820[e]	0.0		0.000	
					17760[f]				
					17670[g]				
					17265[h]				
		1	17992.0	17908.8	18264[e]	83.2	1.49282	1.488	0.00482
					18450[f]				
					18305[g]				
					18327[h]				
		2	19710.4	19710.4	20049[e]	0.0	1.5	1.501	0.001
					20251[f]				
					19886[g]				
					19895[h]				
	^{1}P	1	25068.2	25068.2	24316[e]	0.0	1.035	1.013	0.022
					25967[f]				
					25597[g]				
$4f^{13}5d6s^2$	(7/2,3/2)	2	23188.5	23339.9	24806[e]	151.4	1.45	1.461	0.011
					25028[g]				
					24831[h]				
		5	25859.6	25852.5	26343[g]	7.1	1.04	1.023	0.017
		3	27445.6	27309.2	27064[e]	136.4	1.22	1.223	0.003
					27431[g]				
					27185[h]				
		4	28184.5	28150.4	28013[g]	34.1	1.14	1.137	0.003
	(7/2,5/2)	6	27314.9	27035.7	27205[g]	279.2	1.16	1.167	0.007
		2	28196.0	28358.0	27354[g]	162.0	1.02	1.023	0.003
		1	28857.0	28857.8	29380[e]	0.8	1.2635	1.272	0.0085
					30071[g]				
					29610[h]				
		4	29775.0	29885.2	28975[g]	110.2	1.09	1.091	0.001
		3	30207.4	30215.0	29133[g]	7.6	1.08	1.091	0.011
		5	30524.7	30556.1	29172[g]	31.4	1.18	1.176	0.004
	(5/2,5/2)	0	36047.2				0.000		
		1		39156.9				0.768	
		5		39660.1				1.035	
		2		40499.2				0.918	
		3		41784.5	40189.34[b]			0.979	
		4		42349.4				1.045	

Table 4. Continued.

	(5/2,3/2)	4		37291.3	35882.92[b]			0.827	
		2		38593.0				0.933	
		1		41961.5				0.96	
		3		40740.8	39053.09[b]			0.874	
$4f^{14}6s7p$	^{3}P	0	38090.7	38067.6	38504[f]	23.1		0.000	
		1	38174.2	38204.9	38572[f]	30.7	1.14	1.497	0.357
		2	38552.0	38544.6	38962[f]	7.4	1.5	1.501	0.001
	^{1}P	1	40564.0	40563.7	39311[f]	0.3	1.01	1.005	0.005
$4f^{14}6s5f$	^{1}F	3		42680.6	43297.51[a]			1.045	
	^{3}F	4		42652.3	43319.48[b]			1.251	
		3		43456.6	43254.78[a]			1.039	
		2	43433.9	43433.8		0.0	0.68	0.666	0.014
$4f^{14}6s8p$	^{3}P	0	43614.3	43616.9		2.6		0.000	
		1	43659.4	43656.4		3.0	1.48	1.469	0.011
		2	43805.7	43805.9		0.2	1.49	1.501	0.011
	^{1}P	1	44017.6	44017.8		0.2	1	1.033	0.033
$4f^{14}6s11p$	(1/2,1/2)	0		48192.7				0.000	
		1	48212.1	48198.6		13.5		1.49	
	(1/2,3/2)	2		48215.5				1.501	
		1	48258.5	48271.9		13.4		1.011	
$4f^{14}6s6f$	^{1}F	3		45143.7	45858.57[b]		1.01[b]	1.000	
	^{3}F	2	45956.3	45956.3		0.0	0.72	0.666	0.054
		3		46169.9	45938.90[b]			1.084	
		4		46360.1	45695.89[b]		1.22[b]	1.251	
$4f^{14}6s9p$	^{3}P	0	46082.2	46068.5		13.7		0.000	
		1	46078.9	46095.4		16.5	1.34	1.481	0.141
		2	46184.2	46182.0		2.2	1.5	1.501	0.001
	^{1}P	1	46370.3	46369.6		0.7	1.07	1.02	0.05

[a]Ref. [35], [b]Ref. [36], [c]Ref. [37], [d]Ref. [38], [e]Ref. [32], [f]Ref. [39], [g]Ref. [40], [h]Ref. [41].

Table 5. Yb I even-parity excited state fitting calculated energy levels (in unit cm^{-1}), Landé g-factor values, where ΔE (in unit cm^{-1}) and Δg are absolute errors between experimental and theoretical energies and Landé g-factors, respectively.

			Level				Landé g		
Configuration	Term	J	Expt.[12]	This work	Other work	ΔE	Expt.[12]	This work	Δg
$4f^{14}6s^2$	^{1}S	0	0.0	0.0	0[g]	0.0		0.000	
$4f^{14}5d6s$	^{3}D	1	24489.1	24489.1	27622[e]	0.0	0.5	0.499	0.001
					25108[f]				
					27568[g]				
					27584[h]				
		2	24751.9	24751.9	27632[e]	0.0	1.16	1.164	0.004
					25368[f]				
					27217[g]				
					27678[h]				
		3	25270.9	25275.2	27812[e]	4.3	1.34	1.334	0.006
					25891[f]				
					27747[g]				
					27763[h]				
	^{1}D	2	27677.6	27677.7	28125[e]	0.1	1.01	1.003	0.007
					28353[f]				
					28071[g]				
					28156[h]				
$4f^{14}6s7s$	^{3}S	1	32694.7	32694.7	33092[f]	0.0	2.01	2.002	0.008
	^{1}S	0	34350.6	34350.6	34755[f]	0.0		0.000	
$4f^{14}6s6d$	^{3}D	1	39808.7	39807.6		1.1	0.5	0.499	0.001
		2	39838.0	39839.3		1.3	1.16	1.143	0.017
		3	39966.1	39966.2		0.1	1.33	1.334	0.004
	^{1}D	2	40061.5	40061.3		0.2	1.03	1.024	0.006
$4f^{13}6s^26p$	(7/2,1/2)	3	32065.3	32069.0		3.7	1.23	1.259	0.029
		4	32273.6	32280.3		6.7		1.063	

Table 5. Continued.

	(7/2,3/2)	5	35178.8	35134.8		44.0		1.200	
		2	35197.0	35205.4		8.4	1.05	1.057	0.007
		3	35807.5	35789.2		18.3	1.08	1.088	0.008
		4	36061.0	36104.5		43.5		1.200	
	(5/2,1/2)	3		42343.4	42413.58[d]			0.799	
		2	42531.9	42521.9		10.0		0.965	
	(5/2,3/2)	3		46218.5	46236.27[b]			1.021	
		1	44834.6	44877.6		43.0	0.66	0.499	0.161
		4	45497.6	45473.3		24.3		1.037	
		2	45931.9	45905.1		26.8		0.811	
$4f^{14}6s8s$	^{3}S	1	41615.0	41615.0		0.0	2.02	2.002	0.018
	^{1}S	0	41939.9	41939.9		0.0		0.000	
$4f^{14}6s7d$	^{3}D	1	44311.4	44302.3		9.1		0.499	
		2	44313.1	44314.1		1.0		1.157	
		3	44357.6	44345.3		12.3	1.32	1.334	0.014
	^{1}D	2	44380.8	44401.1		20.3	1.1	1.010	0.09
$4f^{14}6p^2$	^{3}P	0	42436.9	42447.5		10.6		0.000	
		1	43805.4	43763.7		41.7	1.47	1.501	0.031
		2	44760.4	44806.1		45.7	1.34	1.436	0.096
	^{1}D	2	47821.8	47819.8		2.0	1.04	1.065	0.025
	^{1}S	0		49935.2	48344.42[c]			0.000	
$4f^{14}6s8d$	^{3}D	1	46445.0	46451.6		6.6	0.49	0.499	0.009
		2	46467.7	46459.7		8.0	1.12	1.156	0.036
		3	46480.7	46482.0		1.3	1.35	1.334	0.016
	^{1}D	2		46519.5	46405.45[c]			1.011	
$4f^{14}5d^2$	^{3}F	2	47634.4	47634.4		0.0		0.668	
		3	47860.2	47860.3		0.1	1.02	1.084	0.064
		4		48130.4				1.25	
	^{1}D	2		50036.9				1.048	
	^{3}P	0		50404.0				0.000	
		1		50492.5				1.501	
		2		50697.6				1.45	
	^{1}G	4		51319.5				1.000	
	^{4}S	0		56283.8				0.000	

[a]Ref. [35], [b]Ref. [36], [c]Ref. [37], [d]Ref. [38], [e]Ref. [32], [f]Ref. [39], [g]Ref. [40], [h]Ref. [41].

4. Conclusion

For the Yb I atom, we proposed and used the HFR theory in combination with the RR algorithm to calculate 94 energy levels, including the even-parity configurations $4f^{14}6s^2 + 4f^{14}5d6s + 4f^{14}6s7s + \cdots$ and the odd-parity configurations $4f^{14}6s6p + 4f^{13}5d6s^2 + 4f^{14}6s7p + \cdots$. There are 66 experimental energy levels added, and these experimental energy levels fully consider the electron correlation effects. For configurations contain 4f closed-shell electrons, the absolute error between the calculated energy level and the NIST experimental value is less than 50 cm^{-1}, and the relative is less than 0.1%. Among them, the $4f^{13}5d6s^2$ configuration has an average absolute error less than 300 cm^{-1} and a relative error less than 1.0%. Moreover, reliable predicted data are made for energy levels that are experimentally unattainable. This work demonstrates that the ridge regression model has a stronger generalization ability when the amount of data is insufficient, and the predicted data is more reliable. It is expected that a more accurate machine learning model will be applied to the atomic structure calculation.

Acknowledgment

Project supported by the Fundamental Research Funds for the Central Universities (Grant No. 10822041A2038).

References

[1] Ning Y, Jin G Q, Wang M X, Gao S and Zhang J L 2022 *Curr. Opin. Chem. Biol.* **66** 102097
[2] Bünzli J C G 2010 *Chem. Rev.* **110** 2729
[3] Ludlow A D, Boyd M M, Ye J, Peik E and Schmidt P O 2015 *Rev. Mod. Phys.* **87** 637
[4] Eshkabilov N B, Kurbaniyazov A S and Haidarov S R 2022 *Russ. Phys. J.* **64** 1872
[5] Heugel S, Fischer M, Elman V, Maiwald R, Sondermann M and Leuchs G 2016 *J. Phys. B: At. Mol. Opt.* **49** 015002
[6] Sahoo A C, Mandal P K, Mukherjee J, Dev V and Shah M L 2021 *J. Quantum Spectrosc. Radiat. Transfer* **276** 107944
[7] Xu Y W, Shen L and Dai C J 2018 *Mod. Phys. Lett. B* **32** 1850190
[8] Galindo-Uribarri A, Liu Y, Romero Romero E and Stracener D W 2021 *Sci. Rep.* **11** 23432
[9] Furmann B and Stefanska D 2014 *Phys. Scr.* **89** 095402
[10] Kneip N, Düllmann C E, Gadelshin V, Heinke R, Mokry C, Raeder S, Runke J, Studer D, Trautmann N, Weber F and Wendt K 2020 *Hyperfine Interact.* **241** 45
[11] Block M 2019 *Radiochimica Acta* **107** 821

[12] Kramida A, Ralchenko Y, Reader J and NIST ASD Team 2021 *NIST Atomic Spectra Database* (version 5.9), [Online]. Available: https://physics.nist.gov/asd [Sat Jul 09 2022]. National Institute of Standards and Technology, Gaithersburg, MD

[13] Cowan R D 1968 *J. Opt. Soc. Am.* **58** 808

[14] Froese Fischer C, Tachiev G, Gaigalas G and Godefroid M R 2007 *Comput. Phys. Commun.* **176** 559

[15] Hibbert A 1975 *Comput. Phys. Commun.* **9** 141

[16] Jönsson P, Gaigalas G, Bieroń J, Fischer C F and Grant I 2013 *Comput. Phys. Commun.* **184** 2197

[17] Gu M F 2008 *Can. J. Phys.* **86** 675

[18] Desclaux J P 1975 *Comput. Phys. Commun.* **9** 31

[19] Indelicato P, Gorceix O and Desclaux J P 1987 *J. Phys. B: At. Mol. Opt.* **20** 651

[20] Indelicato P and Desclaux J P 1990 *Phys. Rev. A* **42** 5139

[21] Raassen A J J and Uylings P H M 1996 *Phys. Scr.* **T65** 84

[22] Kramida A 2019 *Atoms* **7** 64

[23] Liggins F S, Pickering J C, Nave G, Kramida A, Gamrath S and Quinet P 2021 *Astrophys. J.* **907** 69

[24] Saloman E B and Kramida A 2017 *Astrophys. J., Suppl. Ser.* **231** 18

[25] Saloman E B and Kramida A 2017 *Astrophys. J., Suppl. Ser.* **231** 19

[26] Chikh A, Deghiche D, Meftah A, Tchang-Brillet W Ü L, Wyart J F, Balança C, Champion N and Blaess C 2021 *J. Quantum Spectrosc. Radiat. Transfer* **272** 107796

[27] Hoyt C W, Barber Z W, Oates C W, Fortier T M, Diddams S A and Hollberg L 2005 *Phys. Rev. Lett.* **95** 083003

[28] King A S 1931 *Astrophys. J.* **74** 328

[29] Meggers W F 1942 *Rev. Mod. Phys.* **14** 96

[30] Niyaz F, Nunkaew J and Gallagher T F 2019 *Phys. Rev. A* **99** 042507

[31] Ternovsky V B, Kuznetsova A A, Ternovsky E V, Mironenko D A and Smirnov A V 2018 *J. Phys.: Conf. Ser.* **1136** 012010

[32] Dzuba V A, Flambaum V V and Kozlov M G 2019 *Phys. Rev. A* **99** 032501

[33] Hoerl A E and Kennard R W 1970 *Technometrics* **12** 55

[34] Cowan R D 1981 *The Theory of Atomic Structure and Spectra* (University of California Press)

[35] Bowers C J, Budker D, Commins E D, DeMille D, Freedman S J, Nguyen A T, Shang S Q and Zolotorev M 1996 *Phys. Rev. A* **53** 3103

[36] Wyart J F and Camus P 1979 *Phys. Scr.* **20** 43

[37] Baumann M, Braun M and Maier J 1987 *Z. Phys. D: At., Mol. Clusters* **6** 275

[38] Meggers W and Tech J 1978 *J. Res. Natl. Bur. Stand.* **83** 13

[39] Porsev S G, Safronova M S, Derevianko A and Clark C W 2014 *Phys. Rev. A* **89** 012711

[40] Dzuba V A, Berengut J C, Harabati C and Flambaum V V 2017 *Phys. Rev. A* **95** 012503

[41] Dzuba V A, Flambaum V V and Schiller S 2018 *Phys. Rev. A* **98** 022501