Spectral baseline estimation using penalized least squares with weights derived from the Bayesian method

Qian Wang[1,2] · Xin-Liang Yan[1] · Xiang-Cheng Chen[3] · Peng Shuai[1] · Meng Wang[1,2] · Yu-Hu Zhang[1,2]

Abstract The penalized least squares (PLS) method with appropriate weights has proved to be a successful baseline estimation method for various spectral analyses. It can extract the baseline from the spectrum while retaining the signal peaks in the presence of random noise. The algorithm is implemented by iterating over the weights of the data points. In this study, we propose a new approach for assigning weights based on the Bayesian rule. The proposed method provides a self-consistent weighting formula and performs well, particularly for baselines with different curvature components. This method was applied to analyze Schottky spectra obtained in ^{86}Kr projectile fragmentation measurements in the experimental Cooler Storage Ring (CSRe) at Lanzhou. It provides an accurate and reliable storage lifetime with a smaller error bar than existing PLS methods. It is also a universal baseline-subtraction algorithm that can be used for spectrum-related experiments, such as precision nuclear mass and lifetime measurements in storage rings.

✉ Qian Wang
wangqian2016@impcas.ac.cn

1   Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou 730000, China

2   University of Chinese Academy of Sciences, Beijing 100049, China

3   Van Swinderen Institute, University of Groningen, Groningen 9747 AA, The Netherlands

1 Introduction

A typical one-dimensional spectrum obtained from a spectroscopy experiment is an array that encodes intensity information at a continuous equidistant interval. This can be assumed to be the superposition of a slowly varying baseline, random noise, and a few narrow signal peaks. These signal peaks directly or indirectly reveal the physical and chemical properties of the measured objects. To correctly determine the peak position, intensity, and area, many methods have been developed to eliminate the influence of the baseline and noise, which are commonly observed in spectra. The main goal of these methods is to subtract the baseline from the spectrum while retaining the signal peaks.

Traditional methods for estimating baselines are based on simple or modified polynomial functions that fit the data. Even with automatic optimization in linear programming [1, 2], the performance of the fitting procedure is highly dependent on the knowledge of the baseline, i.e., both the semi-empirical function and initial guess of its fitting coefficients. Its accuracy depends on user experience [3]. As the degrees of freedom of nonlinear fitting increase, fitting instability may become a fatal problem with this method [4].

Modern methods avoid any assumptions about the baseline-specific functional form by using digital filtering [5, 6] and clipping [7]. For example, the rolling ball method [5] operates as a frequency-differentiated nonlinear digital filter by gradually increasing the diameter of the

ball, rolling it below the spectrum, and marking the baseline as the locus of the ball at each point. Peak distortion occasionally occurs after baseline correction and its effect is sometimes not negligible. Another example is the wavelet transform method, which is effective in suppressing random noise in data [6]. This method assumes that the baseline is properly separated from the signal in the transformed domain. However, real-world signals do not always fulfill this assumption. The third example is the sensitive nonlinear iterative peak clipping method, which is included in CERN-ROOT software packages [7, 8]. This method is primarily used in the baseline correction of the proton-induced X-ray emission (PIXE) spectrum [9]. The algorithm compares the values of the data points with a local baseline approximation based on binomial expansions, iteratively stripping sharp peaks and preserving the baseline. One limitation of this approach is its low correction efficiency in regions with relatively large baseline slopes.

In the past two decades, penalized least squares (PLS) baseline estimation has been widely used in the analysis of various types of spectra to overcome the drawbacks mentioned above. The concept of PLS was first proposed in 1922 by Whittaker in his paper on constructing effective smoothers for data [10]. In the 1990s, a similar procedure known as the Hodrick–Prescott filter became popular in the econometric literature [11]. In 2003, Eilers revived the algorithm [12, 13] and applied it to peak alignment and baseline correction [14]. The core of the algorithm is to balance two conflicting goals: fidelity to the data and smoothness of the recommended baseline estimates. Using carefully constructed weights for each data point, the algorithm can effectively extract the baseline shape from the data [14–16]. Owing to the efficient processing of sparse matrices by modern computers, PLS-based algorithms are fast and flexible in terms of implementation. These algorithms do not require knowledge of the specific functions of the baseline and are capable of dealing with complicated baseline shapes. However, the weights of existing methods are empirically determined and may be inconsistent under different conditions, leading to unreasonable results.

In this study, we derived a self-consistent function of weights that adheres to the Bayesian rule for a PLS-based baseline estimation algorithm. In the subsequent sections, the theory of the algorithm and its implementation procedure are introduced. The proposed method is reliable and even more effective than existing PLS-based algorithms in the analysis of synthetic and real Schottky experimental data. In particular, this method stands out in estimating rapidly changing baselines in the presence of signal peaks and strong noise.

2 Theory

2.1 PLS algorithm

The concept of PLS originated from the Whittaker smoother [10], which was initially designed to smooth one-dimensional equally spaced data. Consider a data sequence $\mathbf{y}$ of length N, sampled at equal time intervals. The algorithm smooths a noisy sequence $\mathbf{y}$ into a smooth sequence $\mathbf{z}$ by minimizing the following objective function:

$$S(\mathbf{z}) = (\mathbf{y} - \mathbf{z})^{\mathrm{T}}(\mathbf{y} - \mathbf{z}) + \lambda \mathbf{z}^{\mathrm{T}} \mathbf{D}^{\mathrm{T}} \mathbf{D} \mathbf{z}, \tag{1}$$

where $\mathbf{D}$ is an mth-order difference matrix with $N - m$ rows and N columns. In this study, the discussion is limited to the 2nd-order difference ($m = 2$). Thus, $\mathbf{D}$ is expressed as

$$\mathbf{D} = \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{pmatrix}. \tag{2}$$

The first term in Eq. (1) represents the fidelity of $\mathbf{z}$ to data $\mathbf{y}$, whereas the second term represents the smoothness of $\mathbf{z}$. The balance between fidelity and smoothness is controlled by the parameter λ. A larger λ leads to a smoother $\mathbf{z}$.

In a typical spectrum, some data points have only baseline and noise components, whereas the rest of the points contain additional signal components. For baseline estimation, it is optimal to obtain a slowly varying baseline if there is a way to distinguish between both cases by considering only the former and ignoring the latter. Therefore, combining the Whittaker smoother with the concept of asymmetric least squares [17], Eilers and Boelens [14] first introduced a weight for each point to obtain a smooth baseline. Let $\mathbf{W}$ be a diagonal matrix with a weight array $\mathbf{w}$ of data points on its diagonal. Thus, the objective function in Eq. (1) becomes

$$S(\mathbf{z}) = (\mathbf{y} - \mathbf{z})^{\mathrm{T}} \mathbf{W} (\mathbf{y} - \mathbf{z}) + \lambda \mathbf{z}^{\mathrm{T}} \mathbf{D}^{\mathrm{T}} \mathbf{D} \mathbf{z}. \tag{3}$$

Setting the partial derivatives

$$\frac{\partial S}{\partial \mathbf{z}^{\mathrm{T}}} = -2\mathbf{W}(\mathbf{y} - \mathbf{z}) + 2\lambda \mathbf{D}^{\mathrm{T}} \mathbf{D} \mathbf{z} \tag{4}$$

to zero, we obtain the solution for minimizing Eq. (3):

$$\mathbf{z} = (\mathbf{W} + \lambda \mathbf{D}^{\mathrm{T}} \mathbf{D})^{-1} \mathbf{W} \mathbf{y}. \tag{5}$$

Intuitively, the weight should be 0 if the data point is definitely in the peak region and 1 if the data point does not contain any peak component of the signal. However, the regions containing no signal are difficult to determine because of the presence of noise and peaks. Thus, several

methods have been proposed using iterative algorithms to circumvent peak finding.

A pioneering method is the asymmetric least squares (AsLS) method developed by Eilers and Boelens [14]. In this method, an asymmetry parameter p, recommended to be set between 0.001 and 0.01, is introduced, and the weights are assigned as

$$w_i = \begin{cases} p & y_i > z_i, \\ 1 - p & y_i \leqslant z_i. \end{cases} \tag{6}$$

The algorithm first sets $\mathbf{w}$ to an array of all 1s and solves Eq. (5) to obtain the initial $\mathbf{z}$. Then, the data points are divided into two zones: positive deviations $y_i > z_i$ and negative residuals $y_i \leqslant z_i$. By applying the asymmetric weight function given by Eq. (6), the positive deviations, which are supposed to contain signals, gain fewer weights than the negative residuals in the next iteration. The baseline, i.e., $\mathbf{z}$, is updated by solving Eq. (5) using the new $\mathbf{w}$. By applying this new $\mathbf{z}$, we can again obtain the updated weight array $\mathbf{w}$. The iteration ends when the weight array $\mathbf{w}$ remains unchanged or the number of iterations reaches the maximum pre-set value. The final updated baseline $\mathbf{z}$ is the result of the estimation method.

Given that the AsLS method only assigns two values to the weights of each point, this assignment is unreasonable for data containing different signal peak heights. To improve the weight assignment, the adaptive iteratively reweighted penalized least squares (airPLS) [15] and asymmetrically reweighted penalized least squares (arPLS) [16] methods were successively developed.

The airPLS algorithm considers the effects of signal peak heights. It uses an exponential function to ensure that the weights are small for large peak heights and vice versa. This method also adapts to two-segment weight assignment. The subtraction $\mathbf{d} = \mathbf{y} - \mathbf{z}$, $\mathbf{d}_-$ is part of $\mathbf{d}$ in the region where $\{d_i | d_i = y_i - z_i < 0, d_i \in \mathbf{d}\}$, and $\mathbf{d}_+$ is part of the region where $\{d_i | d_i = y_i - z_i > 0, d_i \in \mathbf{d}\}$. The weight array $\mathbf{w}$ for each iteration step t in the airPLS method is constructed as follows:

$$w_i = \begin{cases} 0 & y_i \geqslant z_i \\ \exp\left[\dfrac{t(y_i - z_i)}{\sum_{d_j \in \mathbf{d}_-} |d_j|}\right] & y_i < z_i. \end{cases} \tag{7}$$

The step number t in Eq. (7) is introduced to speed up the convergence of the iterations. The iteration of this method stops when the maximum number of iterations or the following termination criterion is fulfilled:

$$\sum_{d_j \in \mathbf{d}_-} |d_j| < 0.001 \times |\mathbf{y}|. \tag{8}$$

Given that both the airPLS and AsLS methods attempt to set the weights of the data points to zero, or very close to

zero, they tend to result in a baseline lower than the accurate baseline. To overcome this problem, the arPLS method is proposed, in which an asymmetric weight array is constructed as follows:

$$w_i = \begin{cases} \dfrac{1}{1 + \exp\left[\dfrac{2}{\sigma_-}((y_i - z_i) - (-m_- + 2\sigma_-))\right]} & y_i \geqslant z_i \\ 1 & y_i \leqslant z_i \end{cases}, \tag{9}$$

where m_- and σ_- are, respectively, the mean and standard deviation of $\mathbf{d}_-$.

The arPLS method performs baseline correction as successfully as the AsLS and airPLS methods without a priori peak detection. Indeed, it obtains a more satisfactory baseline in most cases. Similar to the previous method, the weight assignment is still constructed in different forms for both zones ($\mathbf{d}_-$ and $\mathbf{d}_+$). In addition, it does not consider a specific form of the weights at the points corresponding to different signal peak heights. When dealing with higher and wider peaks in a rapidly changing baseline, the results are no longer reliable. In the next section, we propose a unified method to construct weights for the PLS method by invoking the Bayesian rule. The new algorithm outperforms arPLS when it comes to handling rapidly changing baselines and varying peak heights under different noise levels.

2.2 Bayesian approach of weight assignment

Generally, experimental data, denoted by y_i, consist of three components: baseline z_i, noise ϵ_i, and possible signal s_i:

$$y_i = s_i + z_i + \epsilon_i. \tag{10}$$

The baseline and noise are commonly present throughout the spectrum, whereas signals are only present in certain parts. Therefore, some data points in different spectral regions have signal contributions, whereas others do not. As an analogy to the good-and-bad data model [18], the likelihood $\text{prob}(y_i | \beta, \mathcal{I})$ of the data point i can be considered as a weighted combination of case $\mathbf{B}_i$, which contains no signal but only baseline and noise, and case $\overline{\mathbf{B}}_i$, which contains also a signal, as follows:

$$\text{prob}(y_i | \beta, \mathcal{I}) = \beta\, \text{prob}(y_i | \overline{\mathbf{B}}_i, \mathcal{I}) + (1 - \beta)\, \text{prob}(y_i | \mathbf{B}_i, \mathcal{I}), \tag{11}$$

where $0 < \beta < 1$ and the nuisance parameter β represents the probability that a datum contains a signal component. All relevant information $\mathcal{I}$ includes the nature of the physical situation and knowledge of the experiment. The prior probability of case $\mathbf{B}_i$ is

$$\text{prob}(B_i|\beta, \mathcal{I}) = 1 - \beta. \tag{12}$$

The task is to obtain the appropriate weight using Eq. (5) for each point to achieve a smooth baseline z_i. According to the Bayesian theorem [19], the posterior probability for case B_i containing no signal is

$$\text{prob}(B_i|y_i, \mathcal{I}) = \frac{\text{prob}(y_i|B_i, \beta, \mathcal{I})\,\text{prob}(B_i|\beta, \mathcal{I})}{\text{prob}(y_i|\beta, \mathcal{I})}, \tag{13}$$

where $\text{prob}(y_i|B_i, \beta, \mathcal{I})$ is the likelihood function of the data point i in the absence of a signal.

Starting with the common assumption of Gaussian noise with zero mean and standard deviation σ, the likelihood function of case B_i containing no signal can be approximated by a Gaussian probability density function:

$$\text{prob}(y_i|B_i, z_i, \sigma, \mathcal{I}) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(y_i - z_i)^2}{2\sigma^2}\right]. \tag{14}$$

When the datum contains a signal contribution s_i, the likelihood of $\overline{B}_i$ is given by

$$\text{prob}(y_i|\overline{B}_i, z_i, s_i, \sigma, \mathcal{I}) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(y_i - z_i - s_i)^2}{2\sigma^2}\right]. \tag{15}$$

Because s_i is unknown before the analysis, it must be marginalized by

$$\text{prob}(y_i|\overline{B}_i, z_i, \sigma, \mathcal{I}) = \int_0^\infty ds_i\,\text{prob}(y_i|\overline{B}_i, z_i, s_i, \sigma, \mathcal{I}) \\ \text{prob}(s_i|\mathcal{I}). \tag{16}$$

As a convention, s_i is always positive in the task of baseline correction. Here, a common scale factor μ is introduced to describe the signal; let μ be the mean value of all signal heights in the spectrum. The maximum entropy principle leads to a prior probability of s_i [19] under a Poisson distribution

$$\text{prob}(s_i|\mu, \mathcal{I}) = \frac{1}{\mu}\exp\left[-\frac{s_i}{\mu}\right]. \tag{17}$$

Substituting Eqs. (15) and (17) into Eqs. (16) for s_i, we obtain the likelihood function of a data point containing signal as follows:

$$\text{prob}(y_i|\overline{B}_i, z_i, \sigma, \mu, \mathcal{I}) \\ = \frac{1}{2\mu}\left(1 + \text{erf}\left[\frac{y_i - z_i}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\mu}\right]\right) \\ \exp\left[-\frac{y_i - z_i}{\mu} + \frac{\sigma^2}{2\mu^2}\right]. \tag{18}$$

Therefore, the likelihood function in Eq. (11) can be rewritten as follows:

$$\text{prob}(y_i|z_i, \beta, \sigma, \mu, \mathcal{I}) = \beta\,\text{prob}(y_i|\overline{B}_i, z_i, \sigma, \mu, \mathcal{I}) \\ + (1 - \beta)\,\text{prob}(y_i|B_i, z_i, \sigma, \mu, \mathcal{I}) \\ = \beta\,\frac{\exp\left[-\frac{y_i - z_i}{\mu} + \frac{\sigma^2}{2\mu^2}\right]}{2\mu}\left(1 + \text{erf}\left[\frac{y_i - z_i}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\mu}\right]\right) \\ + (1 - \beta)\,\frac{\exp\left[-\frac{(y_i - z_i)^2}{2\sigma^2}\right]}{\sqrt{2\pi}\sigma}. \tag{19}$$

Finally, by substituting the likelihood expressed by Eq. (19) and the prior probability given by Eq. (12) into Eqs. (13), we obtain the posterior probability for a data point i containing no signal:

$$\text{prob}(B_i|y_i, z_i, \beta, \sigma, \mu, \mathcal{I}) \\ = \frac{\text{prob}(y_i|B_i, z_i, \beta, \sigma, \mu, \mathcal{I})\,(1 - \beta)}{p(y_i|z_i, \beta, \sigma, \mu, \mathcal{I})} \\ = \frac{(1 - \beta)\,\text{prob}(y_i|B_i, z_i, \sigma, \mathcal{I})}{\beta\,\text{prob}(y_i|\overline{B}_i, z_i, \sigma, \mu, \mathcal{I}) + (1 - \beta)\,\text{prob}(y_i|B_i, z_i, \sigma, \mathcal{I})} \\ = \frac{1}{1 + f(y_i, z_i, \beta, \sigma, \mu)},$$

where
$$f(y_i, z_i, \beta, \sigma, \mu) = \frac{\beta}{1 - \beta}\sqrt{\frac{\pi}{2}}\frac{\sigma}{\mu} \\ \left(1 + \text{erf}\left[\frac{y_i - z_i}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\mu}\right]\right)\exp\left[\left(\frac{y_i - z_i}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\mu_i}\right)^2\right]. \tag{20}$$

2.3 Bayesian reweighted penalized least squares algorithm

Thus far, we have obtained the posterior probability of a data point containing no signal, expressed by Eq. (20). In this study, this probability value is used as the weight for each point in Eq. (5). Thus, the weight of each datum i can be expressed in a unified format:

$$w_i = \frac{1}{1 + \frac{\beta}{1-\beta}\sqrt{\frac{\pi}{2}}F(y_i - z_i, \sigma, \mu)},$$

$$\text{where} \quad F(d, \sigma, \mu) = \frac{\sigma}{\mu}\left(1 + \text{erf}\left[\frac{d}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\mu}\right]\right) \qquad (21)$$

$$\exp\left[\left(\frac{d}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\mu}\right)^2\right].$$

In practice, nuisance parameters such as the signal chance β, signal scale μ, and Gaussian noise standard deviation σ for each data point must be properly assigned to use the algorithm. Similar to existing algorithms for PLS-based baseline estimation, we also divide the values of subtraction $\mathbf{d} = \mathbf{y} - \mathbf{z}$ into a positive part $\mathbf{d}_+$ and negative part $\mathbf{d}_-$ in the iterations of our algorithm. Then, an approximation of σ and μ can be obtained using different parts of the subtraction: the standard deviation σ can be estimated from the quadratic mean of $\mathbf{d}_-$, whereas the signal scale μ can be estimated from the mean value of $\mathbf{d}_+$.

As mentioned in the previous section, β reflects the probability of signal contribution in the data. According to Eq. (21), if the signal-to-noise ratio (SNR) μ^2/σ^2 [20] and signal scale μ are constant, the calculated weight w_i can be expressed as a function of the subtraction d_i for different values of β, as shown in Fig. 1. The weight functions w_i corresponding to different β values are clearly different.

Thus, a rational value for β is required. Given that the proportion and distribution of signals in the data are unknown, it is convenient to use the average probability of no signal appearance in the total data set as a uniform β factor assigned to each data point, as suggested by the concept of good-and-bad model [18]. In this study, the average probability of no signals in the total data set can be defined as follows:

$$\langle \text{prob}(B_i|y_i, \sigma, \mathcal{I}) \rangle \equiv \frac{1}{N}\sum_{i=1}^{N}\text{prob}(B_i|y_i, \sigma, \mathcal{I}). \qquad (22)$$

Note that β is updated in each iteration and the iteration terminates when β approaches $1 - \langle \text{prob}(B_i|y_i, \sigma, \mathcal{I}) \rangle$. All in all, the algorithm flow of the Bayesian reweighted penalized least squares (BrPLS) method is shown in Algorithm 1.
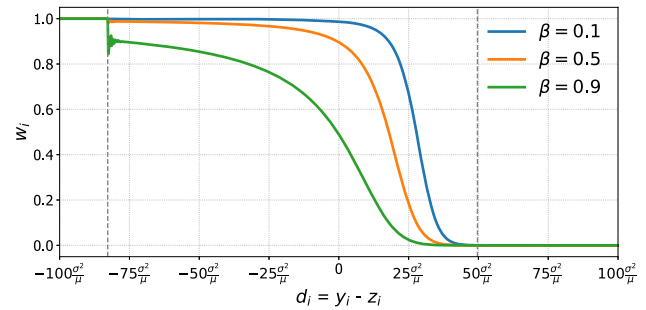

Fig. 1 (Color online) Weight function w_i, defined in Eq. (21), as a function of the subtraction $d_i = y_i - z_i$ for different values of β. The SNR parameter $\mu^2/\sigma^2 = 100$ and signal scale $\mu = 0.1$ are fixed. The region between gray vertical lines highlights a notable change in w_i when β approaches one

---

Algorithm 1 BrPLS method

Input: measured spectra $\mathbf{y}$, smoothness parameter λ,
    termination criteria $ratio$, maximum iteration $maxIter$
Output: estimated baseline $\mathbf{z}$
procedure BRPLS($\mathbf{y}, \lambda, ratio$)
    $\mathbf{z} \leftarrow \mathbf{y}, \beta \leftarrow 0.5$
    $\mathbf{H} = \lambda\mathbf{D}^T\mathbf{D}$ with $\mathbf{D}$ in Eq. (2)
    $\mathbf{w} = [1, 1, \cdots, 1]$
    for $i = 0$ to $maxIter$ **do**
        for $i = 0$ to $maxIter$ **do**
            make $\mathbf{W}$ a diagonal matrix as $W_{i,i} = w_i$
            $\mathbf{z}' = (\mathbf{W} + \mathbf{H})^{-1}\mathbf{W}\mathbf{y}$
            $\mathbf{d} = \mathbf{y} - \mathbf{z}'$
            $\mu = $ mean of positive part of $\mathbf{d}$
            $\sigma = $ quadratic mean of negative part of $\mathbf{d}$
            $\mathbf{F} = \frac{\sigma}{\mu}\left(1 + \text{erf}\left(\frac{\mathbf{d}}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\mu}\right)\right)e^{\left(\frac{\mathbf{d}}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\mu}\right)^2}$
            $\mathbf{w} = 1/(1 + \frac{\beta}{1-\beta}\sqrt{\frac{\pi}{2}}\mathbf{F})$
            if $|\mathbf{z} - \mathbf{z}'|/|\mathbf{z}'| < ratio$ **then**
                break
            else
                $\mathbf{z} \leftarrow \mathbf{z}'$
            end if
        end for
        if $|\beta+ \text{ mean of } \mathbf{w} - 1| < ratio$ **then**
            break
        else
            $\beta \leftarrow 1- \text{ mean of } \mathbf{w}$
        end if
    end for
end procedure

---

3 Algorithm performance

In this section, the performance of the BrPLS method is evaluated along with previous PLS-based methods. The advantages and limitations of these algorithms were revealed by their application to the analysis of synthetic spectra with known baselines and real-world Schottky spectra. All performance tests were performed with codes

programmed in Python using the Numpy and Scipy libraries.

3.1 Reliability tests

We used synthetic spectra to test the reliability of different PLS-based baseline-estimation algorithms. Synthetic spectra were generated by adding signal components and random noise to a pre-defined baseline as follows:

$$d(x) = s(x) + b(x) + n(x), \tag{23}$$

where $s(x)$, $b(x)$, $n(x)$, and $d(x)$ denote the signal, baseline, noise, and resulting synthetic data, respectively. The synthetic signals were eight Gaussian peaks whose intensities, positions, and widths are listed in Table 1. Typical baseline patterns in PIXE [21–23] and Schottky spectra [24] contain linear, concave, or convex regions. Therefore, we chose two types of baseline for the reliability tests: linear and sinusoidal. Random noise was generated by using a random number from a Gaussian distribution. The SNR in terms of energy was defined in the absence of a baseline, according to the following equation [20]:

$$SNR = 10 \log_{10}(E_S/E_n), \tag{24}$$

where E_S refers to the expected energy value of the signal and E_n refers to the expected noise energy value in the spectrum. The SNR of the low-noise spectrum was set to 40 dB and that of the high-noise spectrum was set to 20 dB.

The maximum number of iterations was set to 50 for the AsSL, airPLS, arPLS, and BrPLS methods. The early termination criterion for AsSL, arPLS, and BrPLS was a standard deviation of the difference between two estimated baselines in nearby iterations less than 10^{-6}. For airPLS, Eq. (8) was used as the early termination criterion.

To start the tests of these four methods, an appropriate smoothness parameter λ in Eq. (3) should be determined. Recall that a larger λ leads to a smoother estimated baseline; if λ is too large or too small, the estimated baseline

Table 1 Setting parameters for Gaussian peaks in the synthetic spectra

Peak no.	Intensity	Position	Width
1	10	9	0.7
2	2	20	0.3
3	5	22	0.1
4	15	40	0.2
5	3	49	0.2
6	2	52	0.1
7	20	60	0.6
8	14	70	0.5

will not match the curvature of the baseline or will be highly distorted by high peaks. In addition, the sensitivities of the four methods to the same λ are different. To obtain the best λ for each method, the parameter space of λ was scanned from 10^2 to 10^{12} in 10-base logarithmic steps of 0.1. By the end of the iterations, the algorithm reaches a final set of z_i; then, the root mean square error (RMSE) can be calculated as follows:

$$RMSE(\lambda) = \sqrt{\frac{\sum_{i=1}^{N}(y_i - z_i)^2}{N}}. \tag{25}$$

This RMSE expression reflects the quadratic mean of the estimated signal. If the baseline shape is linear, as λ increases, the RMSE curve first increases rapidly and then converges, as shown in Fig. 2a, c. In this case, λ values can be chosen at any location where the RMSE is stable. For example, the stable $RMSE(\lambda)$ region shown in Fig. 2a corresponds to values of λ ranging from $10^{8.1}$ to 10^{12} and that shown in Fig. 2c corresponds to values of λ ranging from 10^8 to 10^{12}. If the baseline contains depression or protrusion regions, the RMSE as a function of λ can become more complicated, as shown in Fig. 2b, d. A suitable λ value is always chosen at the stable plateau of $RMSE(\lambda)$, where the estimated baseline is close to the actual baseline. For the specific case of a sinusoidal baseline, as in Fig. 2b, the optimal values of λ that can be selected range from $10^{6.5}$ to 10^8 for the AsLS and airPLS methods, from $10^{6.5}$ to $10^{10.2}$ for the arPLS method, and from $10^{6.5}$ to $10^{9.9}$ for the BrPLS method. Similarly, for the case shown in Fig. 2d, the optimal λ values that can be selected range from 10^7 to $10^{8.6}$ for the AsLS and airPLS methods, and from $10^{4.5}$ to $10^{8.6}$ for the arPLS and BrPLS methods. Generally, the expected baseline must be as smooth as possible. Therefore, the maximum optional λ was used in the tests for each method, as listed in Table 2.

To quantitatively assess the performance of each method, we introduced the goodness of the baseline estimation factor R^2 [25]:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(b_i - z_i)^2}{\sum_{i=1}^{N}(\langle b \rangle - b_i)^2}, \tag{26}$$

where b_i and $\langle b \rangle$ denote the synthetic baseline and mean of all the baseline values, respectively. Ideally, the estimated baseline values match the synthetic values, i.e., $R^2 = 1$.

The results are presented in Fig. 3 and Table 2. The BrPLS and arPLS methods have R^2 values close to 1, but the R^2 values of the other two methods deviate from 1. The results suggest that the AsSL and airPLS methods are inadequate for spectral-baseline estimation. The differences in R^2 between the BrPLS and arPLS methods are less than 0.5%, indicating no significant differences between
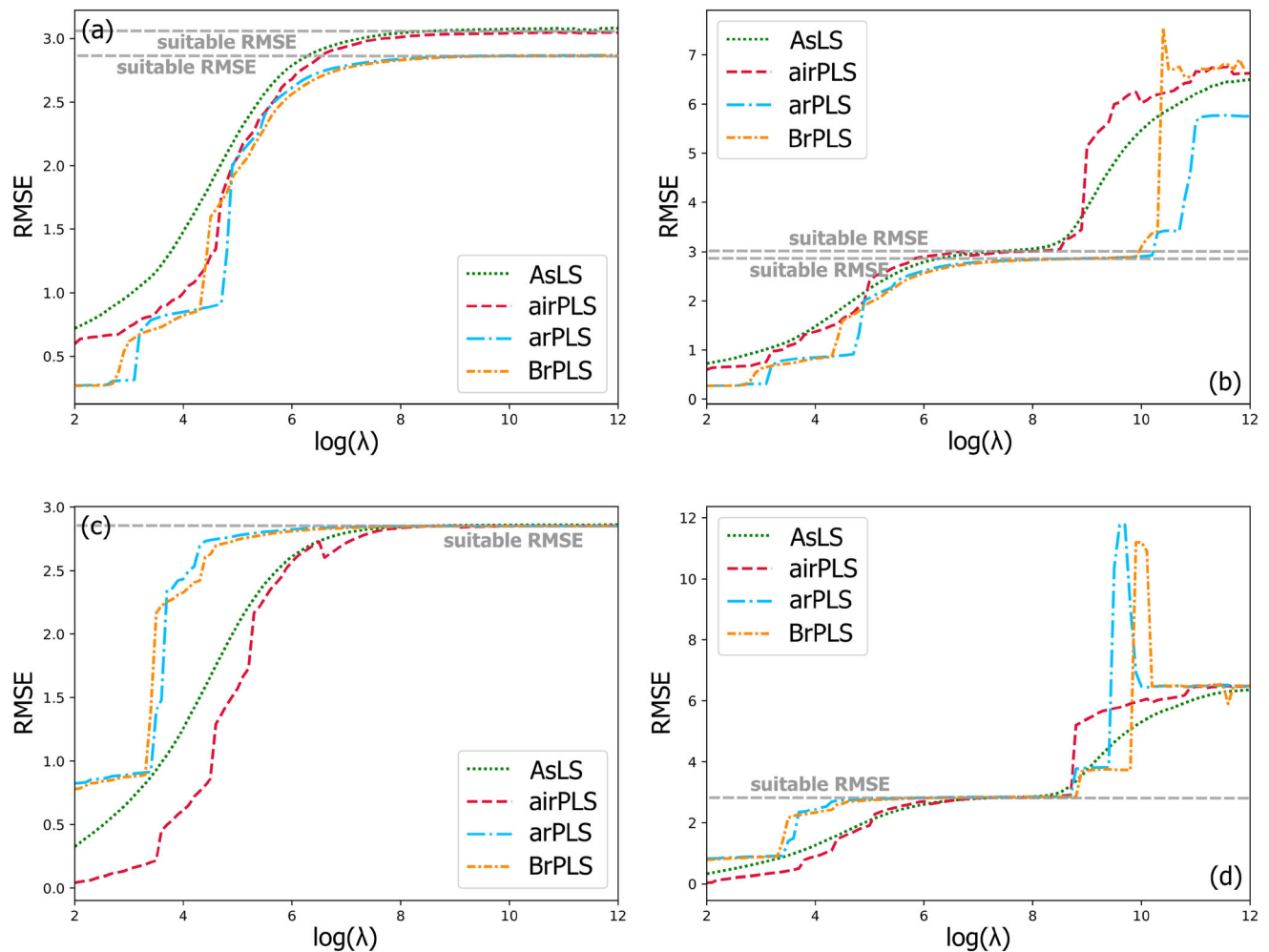
Fig. 2 (Color online) RMSE as a function of the λ parameter calculated from baseline estimation results. **a, c** Depict the case of linear synthetic baseline, while **b, d** depict the case of sinusoidal synthetic baseline. The SNRs used for generating the synthetic spectra were 20 dB in (**a, b**) and 40dB in (**c, d**). The green, red, blue, and orange dash–dot lines represent the results of the AsLS, airPLS, arPLS, and BrPLS methods, respectively. The gray lines indicate the location of the RMSE plateau

Table 2 Results of baseline estimation of the synthetic data shown in Fig. 3

SNR		20 dB			40 dB		
Baseline	Method	$\log(\lambda)$	R^2	RMSE	$\log(\lambda)$	R^2	RMSE
Linear	AsSL	12.0	− 10.8717	3.0775	12.0	0.9418	2.8609
	airPLS	12.0	− 8.7724	3.0539	12.0	0.9827	2.8520
	arPLS	12.0	0.9890	2.8746	12.0	0.9998	2.8521
	BrPLS	12.0	0.9931	2.8731	12.0	0.9998	2.8511
Sinusoidal	AsSL	8.0	− 11.6166	3.0141	8.6	0.6222	2.8135
	airPLS	8.0	− 9.0078	3.0225	8.6	0.8452	2.8374
	arPLS	10.2	0.9791	2.8710	8.6	0.9985	2.8540
	BrPLS	9.9	0.9815	2.8706	8.6	0.9990	2.8518

both methods in these analysis tests. In Fig. 3c, d, the estimated baseline near the data boundary is slightly distorted within the 1-σ error interval of the noise owing to the limitations of the PLS-based methods, where only the 2nd-order difference calculation ($m = 2$) of the baseline was used in Eqs. (1) and (2). In any case, the differences between the estimated and synthetic baselines of both the arPLS and BrPLS methods are within the 1-σ interval of
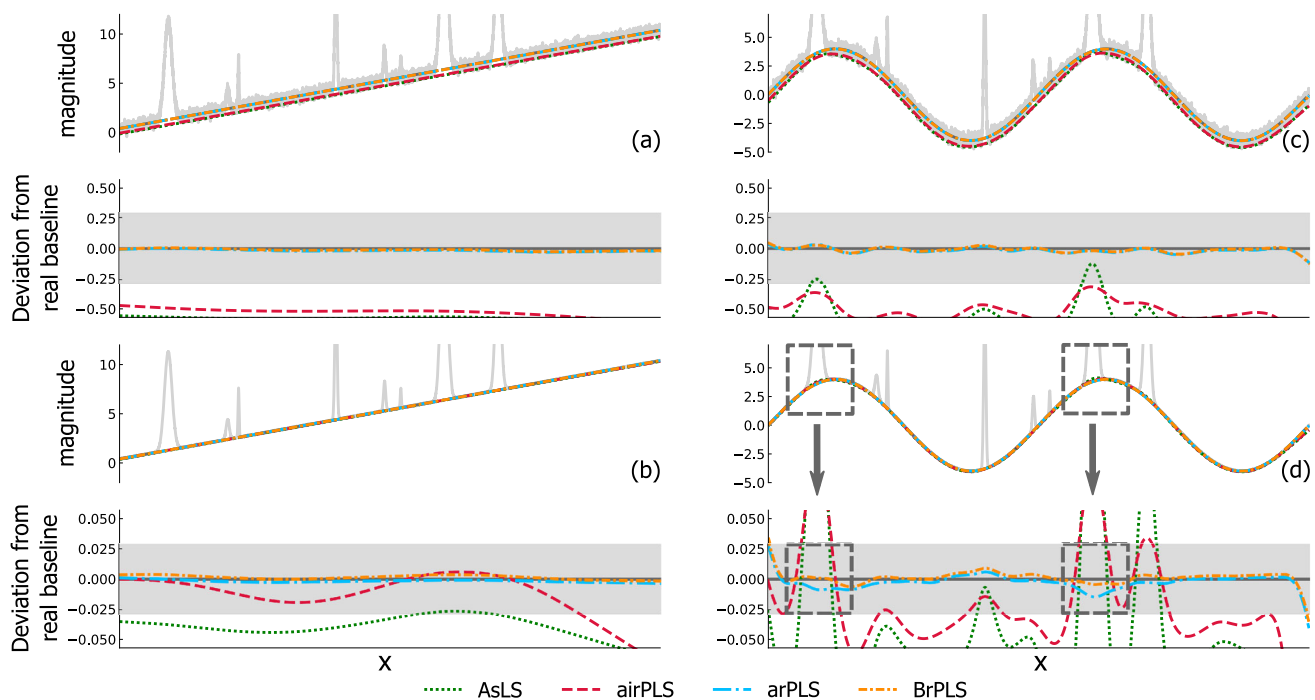
Fig. 3 (Color online) Results of baseline estimation of synthetic spectra. The solid gray lines present the synthetic spectra. The top graph shows the estimated baselines for the four methods and all the original data. The graph below shows the difference between the estimated and synthetic baselines. The interval from $-\sigma$ to σ of the Gaussian noise in the y-axis is highlighted in gray. The green, red, blue, and orange dash–dot lines are estimated baselines using the

AsLS, airPLS, arPLS, and BrPLS methods, respectively. The synthetic spectra depicted in (**a**, **c**) were generated with $SNR = 20$ dB, whereas those shown in (**b**, **d**) were generated with $SNR = 40$ dB. The parts marked with arrows show that the BrPLS method outperforms the arPLS method when dealing with the condition in which signal peaks are located on the baseline protrusion regions

the synthetic Gaussian noise. This means that the estimated baselines of both methods are reliable within the 1-σ error interval. Note also that the BrPLS method has a relatively better ability to handle the signal peaks in the convex regions of the baseline, as indicated by the arrows in Fig. 3d. This will be further evaluated in the next section.

3.2 Comparison between the arPLS and BrPLS methods

According to the results presented in the previous section, the BrPLS and arPLS methods outperform the other two methods regardless of the baseline shape. Moreover, the BrPLS method has a slight advantage over the arPLS method when dealing with cases in which the signal peaks are located in the convex regions of the baseline. In this section, we discuss and compare the advantages and limitations of both algorithms by applying them to the analysis of other synthetic spectra.

The difference between the BrPLS and arPLS methods increases in regions where the baseline has protruding parts. To estimate such a baseline, the most challenging locations are typically those at which the signal peaks lie near the highest point of the convex part and at the

inflection points of the rising and falling edges. Thus, a typical convoluted Landau–Gaussian function was used to generate the synthetic baseline in the synthetic spectrum, and three Gaussian signal peaks were placed in the afore-mentioned parts of the baselines. The intensities and widths of the signal peaks are listed in Table 3. The SNR was set to 40 dB. The maximum number of iterations and early termination criteria for both methods were the same as those in the previous tests. Similarly, the λ values were selected in the same manner as in the previous tests.

The results are presented in Fig. 4 and Table 3. Both methods produce reasonable baseline results. However, when the signal peaks are located at the rising edge of the baseline, the two methods yield unsatisfactory estimations. This is another inherent limitation of PLS-based methods as they strive to balance the smoothness of the baseline and fidelity to the original spectrum. Signal peaks located at the rapidly changing parts of the baseline represent the case in which the two factors are conflicting. As the peak width and height increase, the baseline obtained by the arPLS method further deviates from the actual value when the signal peak is located at the baseline protrusion regions. In contrast, the BrPLS method yields significantly better results, as shown in Fig. 4i. This is because the BrPLS

Table 3 Results of baseline estimations as shown in Fig. 4

Picture label	Peak 1		Peak 2		Peak 3		arPLS			BrPLS		
	Intensity	Width	Intensity	Width	Intensity	Width	$\log(\lambda)$	R^2	RMSE	$\log(\lambda)$	R^2	RMSE
a	3	0.1	5	0.2	10	0.1	7.5	0.9994	0.5309	7.5	0.9997	0.5308
b	3	0.3	5	0.4	10	0.2	7.5	0.9986	0.7609	7.5	0.9987	0.7604
c	3	0.6	5	0.8	10	0.4	7.0	0.9904	01.0795	7.0	0.9910	1.0745
d	6	0.1	10	0.2	20	0.1	7.7	0.9990	1.0617	7.7	0.9993	1.0615
e	6	0.3	10	0.4	20	0.2	7.8	0.9954	1.5221	7.8	0.9964	1.5209
f	6	0.6	10	0.8	20	0.4	7.5	0.9721	2.1592	7.5	0.9771	2.1526
g	12	0.1	20	0.2	40	0.1	8.1	0.9938	2.1235	8.1	0.9963	2.1230
h	12	0.3	20	0.4	40	0.2	8.1	0.9878	3.0452	8.1	0.9903	3.0430
i	12	0.6	20	0.8	40	0.4	8.1	$-$ 1.3622	4.3993	8.1	0.9282	4.3130
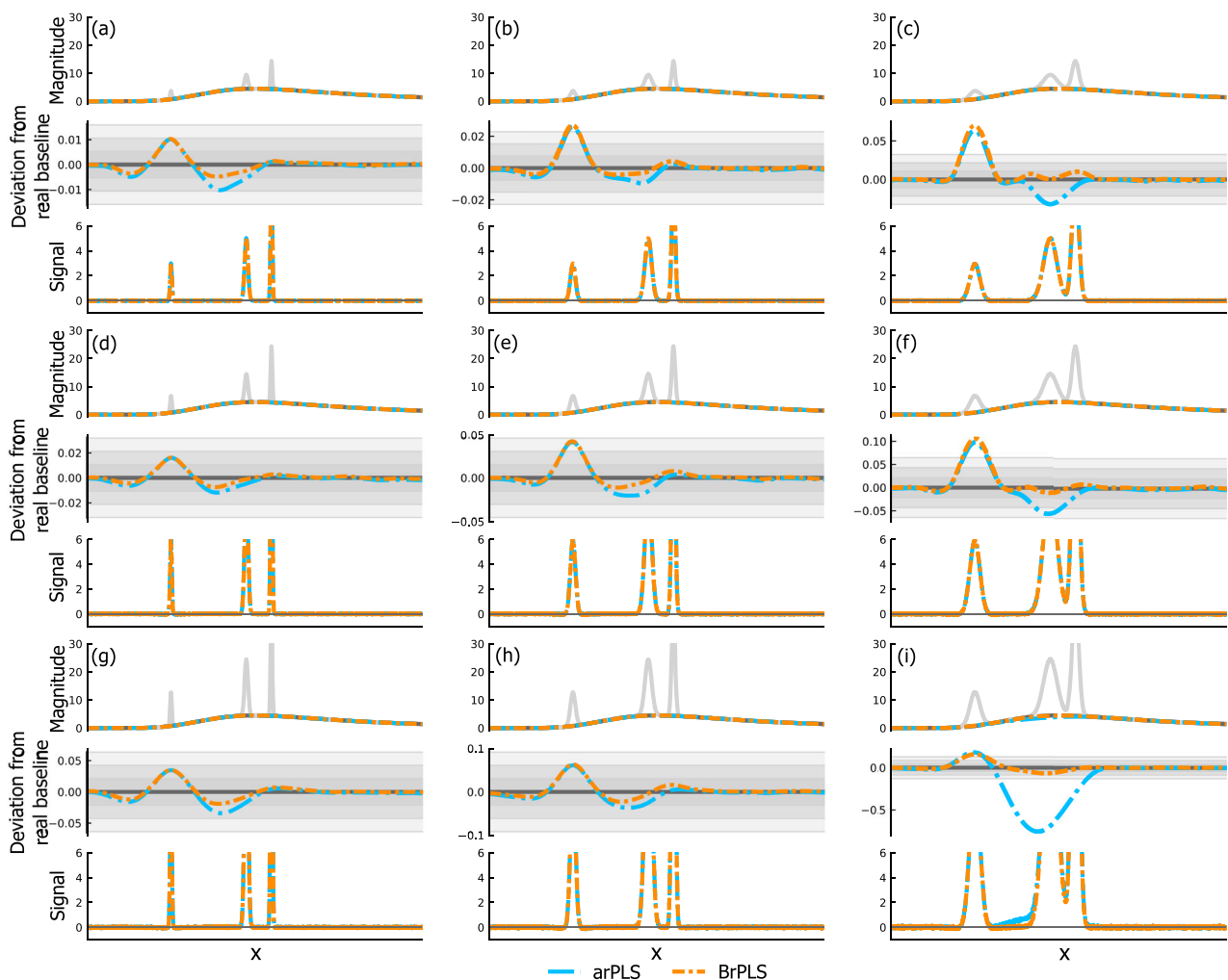
Fig. 4 (Color online) Baseline estimation results using the arPLS (blue dash-dot lines) and BrPLS (dash orange lines) methods. Nine simulation cases, from (**a–i**), share the same convoluted Landau-Gaussian baseline under the same SNR of 40 dB. Additional information for the nine simulations is listed in Table 3. The solid gray lines represent the synthetic spectra. The top graph in each case shows the estimated baselines and original data. The middle graph shows the difference between the estimated and synthetic baselines. The intervals in the y-axis, highlighted in light gray, represent the -3σ-to-3σ intervals of the Gaussian noise. The bottom graph shows the estimated signal magnitudes after baseline subtraction
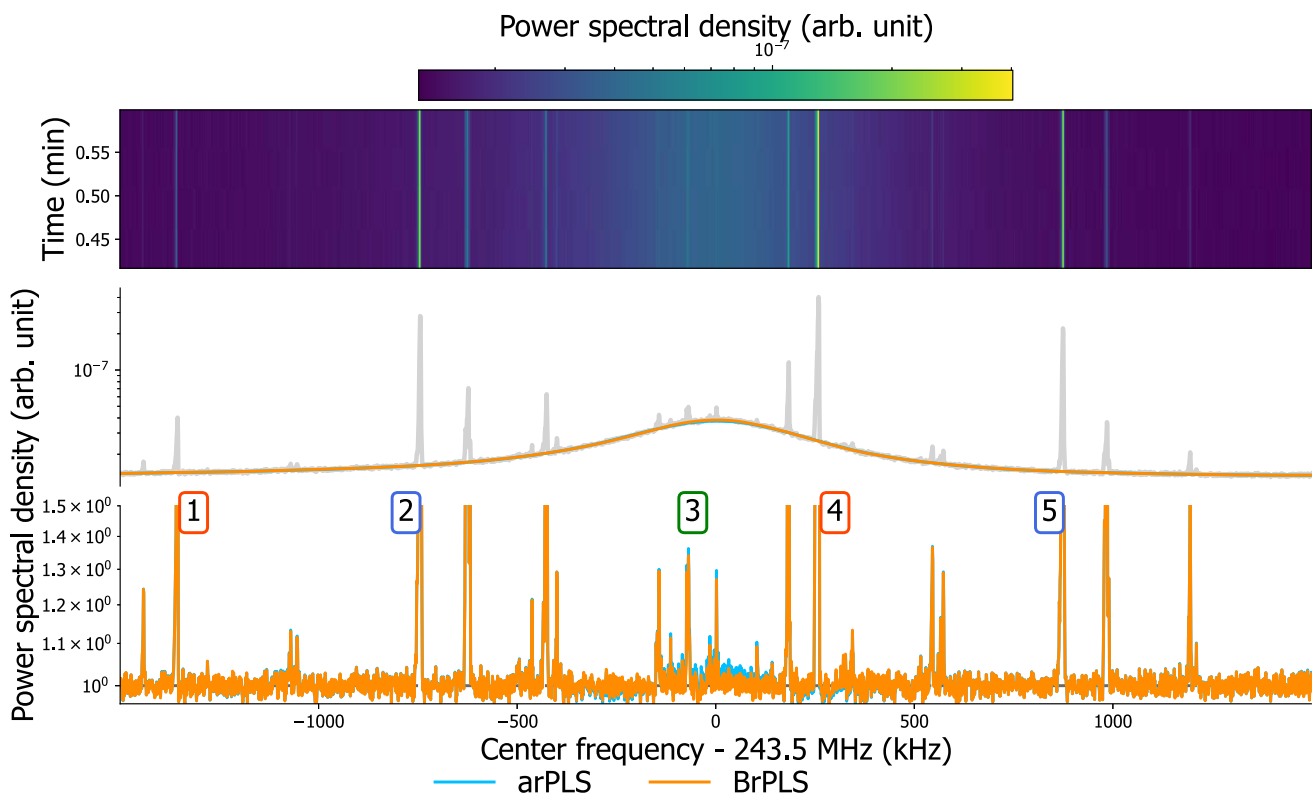
Fig. 5 (Color online) Schottky spectra experimentally measured and corresponding baseline estimation results. The top graph shows the 2D power spectral density in the frequency domain as a function of time after ion injection. The time resolution is 2.73 s. The middle graph shows one frame of spectrum measured at $t = 25$ s. The power spectrum after baseline correction is presented in the bottom graph. Five ion peaks are marked for further discussion. Ion peaks belonging to the same ion species but with different harmonic numbers of the revolution frequency are identified by the same color box: peaks 1 and 4 belong to $^{82}Se^{34+}$, peaks 2 and 5 belong to $^{84}Br^{35+}$, and peak 3 belongs to $^{74}Ga^{31+}$

method considers the heights of different signal peaks and assigns different weight values to the parameter μ in Eq. (21). Conversely, the weight function in Eq. (9) for the arPLS method is a symmetric sigmoid curve that does not consider signal peak characteristics.

3.3 Application to real-world Schottky spectra

Schottky spectroscopy [26] is an important method for beam diagnostics in heavy-ion accelerator facilities. It is routinely used in precision mass and lifetime measurement experiments in nuclear physics studies [27]. The corresponding Schottky spectrum reflects the power density of signals induced by circulating ions in a storage ring. The periodic motion of an ion is represented by the resonance peaks at each harmonic of the revolution frequency in the measured spectrum [28]. Accurate determination of the peak position and area in the power density spectrum is essential for data analysis. The revolution frequency, which is proportional to the mass-to-charge ratio of the ion, can be derived from the peak position. The peak area can be derived from the baseline-corrected spectrum. It is proportional to the number of corresponding ions and can be used to determine the lifetime of the ions in the storage ring.

In the Experimental Storage Ring at Darmstadt [29] and experimental Cooler Storage Ring (CSRe) at Lanzhou [30], ions circulate with frequencies ranging from several hundred kHz to a few MHz. The bandwidth of the measured Schottky spectrum is typically less than 2 MHz [31–33]. The baseline of this type of spectrum can be approximated as a straight line and, in some cases, can be corrected using simple 1st-order polynomial fittings [34, 35]. However, the baseline becomes complicated for the full spectrum, which contains several harmonics of the full momentum acceptance of the storage ring. Owing to complex signal processing in hardware [24], the baseline appearing in this spectrum does not always have a theoretical Lorentzian shape [36]. Different types of baseline distortions may occur during the actual signal processing. Thus, a method that can effectively estimate and correct Lorentzian-type

baselines is required. The newly developed BrPLS method offers a new option in this direction.

The power spectral density of Schottky noise is multiplicative [37] and our Bayesian assumption is based on the additive components in the data, as expressed in Eq. (10). Therefore, a logarithmic transformation of the power densities in the original Schottky spectra was applied first. The processed spectra were then adjusted to non-negative values by subtracting the minimum values. Experimental data were obtained from the measurements of ^{86}Kr projectile fragments [38–40] at CSRe in 2018. Each frame in the Schottky spectrum had a data acquisition time of 2.73 s. The statistical uncertainty of the amplitude at a given frequency point was approximately 1.4% [41]. Particle identification in the spectrum was performed according to [26]. During baseline estimation, steps similar to those of the previously described synthetic data treatment were applied. The parameters and other settings of the arPLS and BrPLS methods were the same as those described in the previous

section. The results of a simple baseline fit with a local 1st-order polynomial (fitting frequency range between $\pm$ 30 kHz of the signal peak center) are also presented for comparison. The peak position and area under the frequency peak were determined by fitting Gaussian functions to the peaks after baseline corrections.

The baseline estimation results for a single frame are shown in Fig. 5. Consistent with the previous section, the BrPLS method could effectively handle the presence of peaks in the baseline protrusion regions. In contrast, the arPLS method underestimated the baseline to the extent of introducing more spurious signal peaks, as indicated by the blue line in Fig. 6. In Fig. 5, five signal peaks are marked for further discussion. Peaks 1 and 4 belong to the same ion species ^{82}Se^{34+} but differ in the harmonic number of the ion revolution frequency. Peak 1 is located in the flat baseline part, whereas peak 4 is located in the middle of the falling edge of the baseline. Peaks 2 and 5 belong to the same ion species ^{84}Br^{35+} and are located in the relatively flat part of
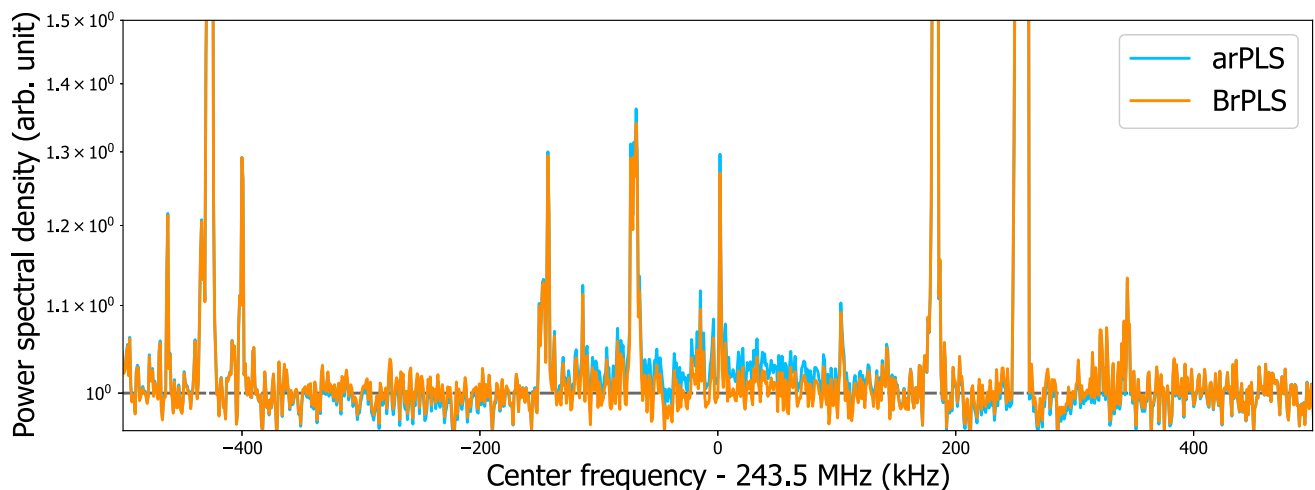

Fig. 6 (Color online) Baseline-corrected spectrum, as shown in the bottom graph of Fig. 5, zoomed into the frequency range of $(-500, 500)$ kHz
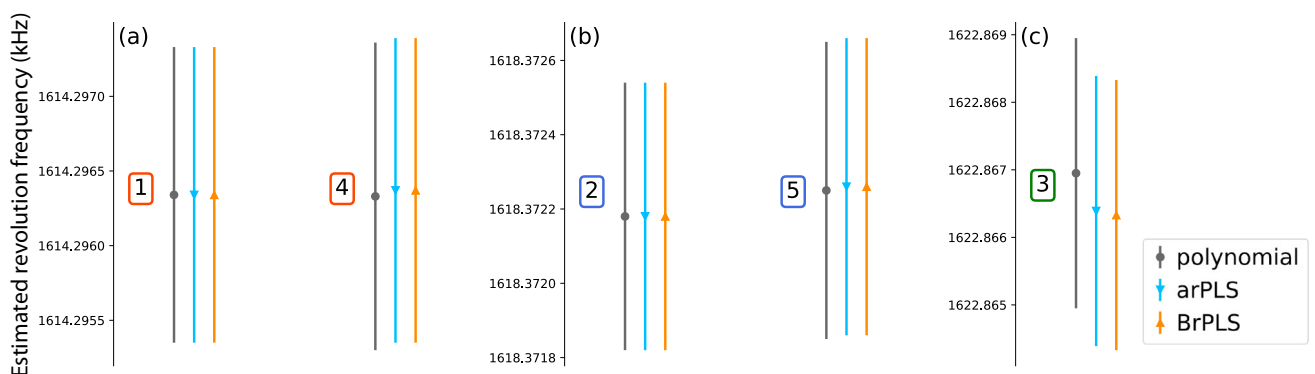

Fig. 7 (Color online) Revolution frequency determined from the baseline-corrected spectrum. The peak positions were extracted from Gaussian peak fitting. The revolution frequencies were normalized by

the corresponding harmonic numbers. The labels of the ion peaks are the same as in Fig. 5
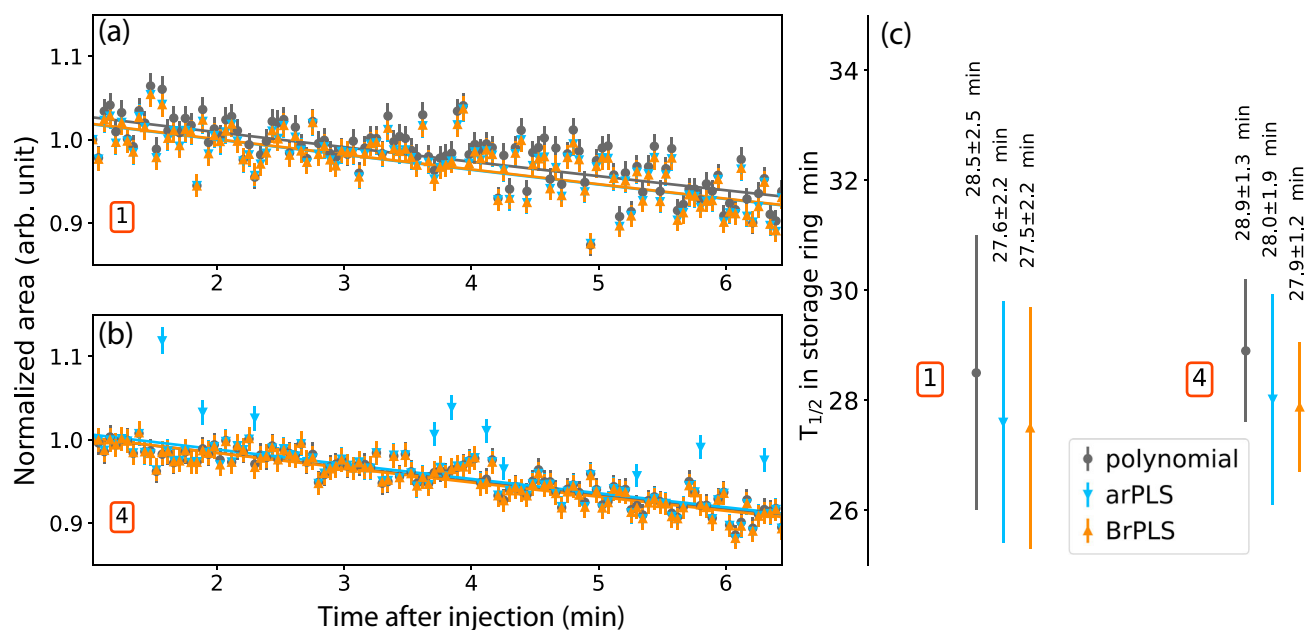
Fig. 8 (Color online) Lifetime estimation results of $^{82}Se^{34+}$ (neutral atom half-life $T_{1/2} = 87.6$ Ey [42]) using data (peaks 1 and 4) in Fig. 5. The left graphs **a**, **b** show peak areas, normalized to the initial area after baseline correction, as a function of the time. The solid lines are exponential-function fitting results. The labels of the ion peaks are the same as those in Fig. 5. The error bars represent statistical uncertainties of the signal areas extracted from the Gaussian fits. The deduced storage half-lives of the corresponding ions in CSRe are shown in graph (**c**)

the baseline. The small ion peak 3 belongs to $^{74}Ga^{31+}$ ions and is located near the top of the baseline-protrusion region. The estimated revolution frequencies were consistent among the three methods within the 1σ error range, as shown in Fig. 7. The central values vary slightly as a function of the peak position and method, particularly when the peak is located in the convex part of the baseline. The peak area as a function of the elapsed time was used to determine the ion storage lifetime after ion injection into the ring. More spectra shown in the top graph of Fig. 5 were analyzed for this task, and the differences in the performance of the three methods became more apparent. Ideally, the ion lifetimes estimated from the ion peaks at different harmonic numbers were expected to be the same. As shown in Fig. 8, the three methods achieved this goal. However, arPLS incurred a relatively larger error for ion peak 4, which is located in the baseline protrusion region. In this region, the ion peak areas obtained by the arPLS method were somewhat overfitted, resulting in an underestimation of the baseline. Some blue points obtained using this method and depicted in Fig. 8b deviate from the overall trend, leading to larger uncertainty in the corresponding lifetime, as shown in Fig. 8c. The overall difference between the results of the polynomial fitting and BrPLS methods is negligible. It is expedient to use a local linear fit to the baseline to estimate the area of individual peaks when the baseline shape of the full spectrum is very complex. However, this method relies heavily on the degree of linearity of the baseline within the selected local

range. Additionally, to estimate the neutral atomic lifetime of the corresponding nuclide of an ion from its storage lifetime in the ring, it is necessary to use other stored ions in the same spectrum as a reference. A local linear fit to individual peaks is challenging to guarantee the consistency of the baseline estimate for all the peaks located in different parts of the spectrum. In contrast, the results obtained by the newly developed BrPLS method are generally more stable and reliable than those of the other two methods.

4 Summary and outlooks

This paper reports the BrPLS method for baseline estimation, which provides a unified and reliable weight assignment based on the Bayesian theorem. This method is an upgrade to the existing PLS method and is implemented as an iteration algorithm. The new method was applied to the treatment of both synthetic and experimental spectra. It outperforms existing PLS-based baseline estimation methods and yields relatively better results, especially when the signal peaks lie on the convex part of the baseline in the presence of noise.

An open-source program implementing the BrPLS method can be downloaded from Github [43], providing an alternative baseline correction method for spectrum analysis. Further developments may include the following: (1) using Poisson noise instead of Gaussian noise (ϵ_i) in

Eq. (10); (2) improving the reliability of the baseline estimations using datum-specific scale factors (μ_i) for each bin instead of an average μ in Eq. (17); (3) similarly, assigning datum specific (β_i) instead of an average β in Eq. (11) to the probability that the datum contains signal components.

References

1. F. Gan, G.H. Ruan, J.Y. Mo, Baseline correction by improved iterative polynomial fitting with automatic threshold. Chemom. Intell. Lab. Syst. **82**(1–2), 59–65 (2006). https://doi.org/10.1016/j.chemolab.2005.08.009

2. Z.M. Zhang, S. Chen, Y.Z. Liang et al., An intelligent background-correction algorithm for highly fluorescent samples in Raman spectroscopy: background-correction algorithm for highly fluorescent samples in Raman spectroscopy. J. Raman Spectrosc. **41**(6), 659–669 (2009). https://doi.org/10.1002/jrs.2500

3. A. Jirasek, G. Schulze, M.M.L. Yu et al., Accuracy and precision of manual baseline determination. Appl. Spectrosc. **58**(12), 1488–1499 (2004). https://doi.org/10.1366/0003702042641236

4. W. von der Linden, V. Dose, L. Padayachee et al., Signal and background separation. Phys. Rev. E **59**(6), 6527–6534 (1999). https://doi.org/10.1103/physreve.59.6527

5. M.A. Kneen, H.J. Annegarn, Algorithm for fitting XRF, SEM and PIXE X-ray spectra backgrounds. Nucl. Instrum. Methods B **109**(110), 209–213 (1996). https://doi.org/10.1016/0168-583X(95)00908-6

6. X.G. Ma, Z.X. Zhang, Application of wavelet transform to background correction in inductively coupled plasma atomic emission spectrometry. Anal. Chim. Acta **485**(2), 233–239 (2003). https://doi.org/10.1016/S0003-2670(03)00395-7

7. M. Morháč, V. Matoušek, Peak clipping algorithms for background estimation in spectroscopic data. Appl. Spectrosc. **62**(1), 91–206 (2008). https://doi.org/10.1366/000370208783412762

8. C.G. Ryan, E. Clayton, W.L. Griffin et al., SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications. Nucl. Instrum. Methods B **34**, 396–402 (2013). https://doi.org/10.1016/0168-583X(88)90063-8

9. J.C. Lindon, G.E. Tranter, D.W. Koppenaal, *Encyclopedia of spectroscopy and spectrometry*, 2nd edn. (Elservier Academic Press, Oxford, 2010)

10. E.T. Whittaker, On a new method of graduation. Proc. Edinb. Math. Soc. **41**, 63–75 (1922). https://doi.org/10.1017/S0013091500077853

11. R.J. Hodrick, E.C. Prescott, Postwar U.S. business cycles: an empirical investigation. J. Money Credit Bank. **29**(1), 1–16 (1997). https://doi.org/10.2307/2953682

12. P.H.C. Eilers, A perfect smoother. Anal. Chem. **75**(14), 3631–3636 (2003). https://doi.org/10.1021/ac034173t

13. P.H.C. Eilers, V. Pesendorfer, R. Bonifacio, in Proceedings of 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Bruges, June (2017)

14. P.H.C. Eilers, Parametric time warping. Anal. Chem. **76**(2), 404–411 (2004). https://doi.org/10.1021/ac034800e

15. Z.M. Zhang, S. Chen, Y.Z. Liang, Baseline correction using adaptive iteratively reweighted penalized least squares. Analyst **135**(5), 1138–1146 (2010). https://doi.org/10.1039/b922045c

16. S.J. Baek, A. Park, Y.J. Anh et al., Baseline correction using asymmetrically reweighted penalized least squares smoothing. Analyst **140**(1), 250–257 (2015). https://doi.org/10.1039/C4AN01061B

17. W.K. Newey, J.L. Powell, Asymmetric least squares estimation and testing. Econometrica **55**(4), 819–847 (1987). https://doi.org/10.2307/1911031

18. D.S. Sivia, in MAXENT96—Proceedings of the Maximum Entropy Conference, Port Elizabeth, ed. by M. Sears, V. Nedeljkovic, N.E. Pendock, S. Sibisi (1996), p. 131

19. D.S. Sivia, J. Skilling, *Data Analysis: A Bayesian Tutorial*, 2nd edn. (Oxford University Press, Oxford, 2006), pp.5–13

20. B. Carlson, P.B. Crilly, *Communication Systems: An Introduction to Signals and Noise in Electrical Communication*, 5th edn. (McGraw-Hill Higher Education, Boston, 2010), pp.423–426

21. G.C.-Y. Chan, in *Encyclopedia of Analytical Science*, (3rd edn) ed. by P. Worsfold, C. Poole, A. Townshend, M. Miró (Elsevier, Amsterdam, 2019), pp. 194–299. https://doi.org/10.1016/B978-0-12-409547-2.14487-7

22. H.W. Wang, G.T. Fan, L.X. Liu et al., Commissioning of laser electron gamma beamline SLEGS at SSRF. Nucl. Sci. Tech. **33**, 87 (2022). https://doi.org/10.1007/s41365-022-01076-0

23. R.M.J. Li, S.K. Liu, S.T. Lin et al., Identification of anomalous fast bulk events in a p-type point-contact germanium detector. Nucl. Sci. Tech. **33**, 57 (2022). https://doi.org/10.1007/s41365-022-01041-x

24. C. Trageser, C. Brandau, C. Kozhuharov et al., A new data acquisition system for Schottky signals in atomic physics experiments at GSI's and FAIR's storage rings. Phys. Scr. **T166**, 014062 (2015). https://doi.org/10.1088/0031-8949/2015/T166/014062

25. D. Chicco, M.J. Warrens, G. Jurman, The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. Peer. J. Comput Sci. (2021). https://doi.org/10.7717/peerj-cs.623

26. Yu.A. Litvinov, H. Geissel, T. Radon et al., Mass measurement of cooled neutron-deficient bismuth projectile fragments with time-resolved Schottky mass spectrometry at the FRS-ESR facility. Nucl. Phys. A **756**(1–2), 3–38 (2005). https://doi.org/10.1016/j.nuclphysa.2005.03.015

27. Yu.A. Litvinov, S. Bishop, K. Blaum et al., Nuclear physics experiments with ion storage rings. Nucl. Instrum. Methods B **317**, 603–616 (2013). https://doi.org/10.1016/j.nimb.2013.07.025

28. S. Chattopadhyay, Some fundamental aspects of fluctuations and coherence in charged-particle beams in storage rings. AIP Conf. Proc. **127**, 467 (1985). https://doi.org/10.1063/1.35175

29. B. Franzke, The heavy ion storage and cooler ring project ESR at GSI. Nucl. Instrum. Methods B **24–25**, 18–25 (1987). https://doi.org/10.1016/0168-583X(87)90583-0

30. W.L. Zhan, J.W. Xia, B.W. Wei et al., HIRFL-CSR project. AIP Conf. Proc. **600**, 175 (2001). https://doi.org/10.1063/1.1435229

31. B. Franzke, H. Geissel, G. Münzenberg, Mass and lifetime measurements of exotic nuclei in storage rings. Mass Spectrom. Rev. **27**, 428–469 (2008). https://doi.org/10.1002/mas.20173

32. B. Bosch, Yu.A. Litvinov, T. Stöhlker, Nuclear physics with unstable ions at storage rings. Prog. Part. Nucl. Phys. **73**, 84–140 (2013). https://doi.org/10.1016/j.ppnp.2013.07.002

33. Yu.A. Litvinov, F. Bosch, Beta decay of highly charged ions. Rep. Prog. Phys. **74**(1), 016301 (2011). https://doi.org/10.1088/0034-4885/74/1/016301

34. Yu.A. Litvinov, Ph.D. thesis, der Justus-Liebig Universität Gießen (2003)

35. X.L. Tu, X.C. Chen, J.T. Zhang et al., First application of combined isochronous and Schottky mass spectrometry: half-lives of fully ionized $^{49}CR^{24+}$ and $^{53}Fe^{26+}$ atoms. Phys. Rev. C **97**, 014321 (2018). https://doi.org/10.1103/PhysRevC.97.014321

36. X.C. Chen, Ph.D. thesis, Ruprecht-Karls-Universität Heidelberg (2015)

37. X.C. Chen, Yu.A. Litvinov, M. Wang et al., Denoising scheme based on singular-value decomposition for one-dimensional spectra and its application in precision storage-ring mass spectrometry. Phys. Rev. E **99**(6), 063320 (2019). https://doi.org/10.1103/PhysRevE.99.063320

38. C.W. Ma, D. Peng, H.L. Wei et al., A Bayesian-neural-network prediction for fragment production in proton induced spallation reaction. Chin. Phys. C **44**(12), 124107 (2020). https://doi.org/10.1088/1674-1137/abb657

39. C.W. Ma, H.L. Wei, X.Q. Liu et al., Nuclear fragments in projectile fragmentation reactions. Prog. Part. Nucl. Phys. **121**, 103911 (2021). https://doi.org/10.1016/j.ppnp.2021.103911

40. C.W. Ma, H.L. Wei, X.X. Chen et al., Precise machine learning models for fragment production in projectile fragmentation reactions using Bayesian neural networks. Chin. Phys. C **46**(7), 074104 (2022). https://doi.org/10.1088/1674-1137/ac5efb

41. B. Schlitt, Ph.D. thesis, Ruprecht-Karls-Universität Heidelberg (1997)

42. F.G. Kondev, M. Wang, W.J. Huang et al., The NUBASE2020 evaluation of nuclear physics properties. Chin. Phys. C **45**(3), 030001 (2021). https://doi.org/10.1088/1674-1137/abddae

43. Q. Wang, baseline estimation program, Github (2021). https://github.com/NanaVan/baseline-estimate. Accessed 8 Nov 2021