

# Rapid deconvolution of low-resolution time-of-flight data using Bayesian inference F SCI

Cite as: J. Chem. Phys. **151**, 244307 (2019); <https://doi.org/10.1063/1.5129343>

Submitted: 27 September 2019 • Accepted: 02 December 2019 • Published Online: 30 December 2019

 Cornelius L. Pieterse,  Michiel B. de Kock,  Wesley D. Robertson, et al.

## COLLECTIONS

F This paper was selected as Featured

SCI This paper was selected as Scilight



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

[New deconvoluting algorithm offers new approach for mass spectrometry data compression](#)

Scilight **2020**, 011103 (2020); <https://doi.org/10.1063/10.0000540>

[Solving the chaotic source separation problem](#)

Scilight **2020**, 061114 (2020); <https://doi.org/10.1063/10.0000783>

[X-ray diffractive imaging of controlled gas-phase molecules: Toward imaging of dynamics in the molecular frame](#)

The Journal of Chemical Physics **152**, 084307 (2020); <https://doi.org/10.1063/1.5133963>



Webinar  
Quantum Material Characterization  
for Streamlined Qubit Development



Register now




# Rapid deconvolution of low-resolution time-of-flight data using Bayesian inference

Cite as: J. Chem. Phys. 151, 244307 (2019); doi: 10.1063/1.5129343

Submitted: 27 September 2019 • Accepted: 2 December 2019 •

Published Online: 30 December 2019



Cornelius L. Pieterse,<sup>1,a)</sup>  Michiel B. de Kock,<sup>1,2</sup>  Wesley D. Robertson,<sup>1</sup>  Hans C. Eggers,<sup>2,3</sup>  
and R. J. Dwayne Miller<sup>1,4,b)</sup> 

## AFFILIATIONS

<sup>1</sup>Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany

<sup>2</sup>Department of Physics, Stellenbosch University, Matieland 7602, South Africa

<sup>3</sup>National Institute for Theoretical Physics, Matieland 7602, South Africa

<sup>4</sup>Departments of Chemistry and Physics, University of Toronto, 80 St. George Street, Ontario, Toronto M5S 3H6, Canada

<sup>a)</sup>Present address: National Physical Laboratory, Hampton Road, Teddington TW11 0LW, United Kingdom.

<sup>b)</sup>Electronic mail: [dwayne.miller@mpsd.mpg.de](mailto:dwayne.miller@mpsd.mpg.de)

## ABSTRACT

The deconvolution of low-resolution time-of-flight data has numerous advantages, including the ability to extract additional information from the experimental data. We augment the well-known Lucy-Richardson deconvolution algorithm using various Bayesian prior distributions and show that a prior of second-differences of the signal outperforms the standard Lucy-Richardson algorithm, accelerating the rate of convergence by more than a factor of four, while preserving the peak amplitude ratios of a similar fraction of the total peaks. A novel stopping criterion and boosting mechanism are implemented to ensure that these methods converge to a similar final entropy and local minima are avoided. Improvement by a factor of two in mass resolution allows more accurate quantification of the spectra. The general method is demonstrated in this paper through the deconvolution of fragmentation peaks of the 2,5-dihydroxybenzoic acid matrix and the benzyltriphenylphosphonium thermometer ion, following femtosecond ultraviolet laser desorption.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5129343>

## I. INTRODUCTION

Mass spectrometry (MS) is practiced in various settings, including commercial applications such as quality control<sup>1</sup> and pharmacokinetics,<sup>2</sup> as well as basic scientific investigations such as proteomics<sup>3</sup> and biological pathway analysis.<sup>4</sup> While the operation of mass spectrometers can be regarded as a routine and high-throughput task, the correct interpretation of the spectra requires a chemistry background and often ample experience. Recent interest in the detection of biomarkers<sup>5–7</sup> further underlines the significance of accurate and thorough interpretation of mass spectrometric data. It has been demonstrated that it is possible to distinguish between healthy and unhealthy domains of mammalian tissue sections by comparison with their respective mass spectrometric data.<sup>8–10</sup> This opens up the possibility of compiling a mass spectrometric database of biomarkers associated with recognized

diseases for the identification of unhealthy tissue.<sup>10–12</sup> The ability to distinguish between these domains should permit the accurate identification of critical boundaries using mass spectrometric imaging techniques.

When investigating an extensive mass range, the large amounts of acquired data present an inescapable dilemma: one must either deal with very large raw datasets or else compress the raw data by averaging. While compression is a simple and adequate solution for many applications, there are cases where it destroys valuable information. A typical example would be pulsed laser beam analyses where it might be of interest to study the mass spectra as a function of depth or shot number.<sup>13–16</sup>

While it has always lurked in the background, the data compression dilemma has not been acute in mass spectrometry so far, largely because investigations and results were either qualitative in nature or couched in simple terms such as one-dimensional

time-independent data. As more complex measurements such as mass spectrometry imaging become feasible, the negative effects of the trade-off between much larger datasets and information loss are increasingly being felt. Each additional variable or differential quantity results in a significant increase in the number of possible outcomes, resulting in low signal-to-noise ratios.

While it is possible to obtain favorable results by the brute force of long acquisition times, this may have unintended consequences. As an example, more accurate measurements could be obtained by the adoption of the minimum focal spot size of the laser beam to maximize resolution in mapping the boundary between two domains. The resulting longer acquisition time, however, is self-defeating because signal intensity drops in time due to the decreasing number of ions available for sampling.<sup>17–20</sup> It is common knowledge that the concentration of the biomarkers of interest is usually orders of magnitude lower than that of the closely related mass species in the surrounding tissue which frequently includes species which contribute to ion suppression and therefore impede detection of these biomarkers.<sup>3,21,22</sup>

In a similar vein, mass spectrometry signals originating from a variety of ion sources are generally not constant in time. These time variabilities introduce correlations and thereby obscure the genuine average mass spectrum. For example, the intensity chromatographs of liquid samples irradiated by laser pulses under atmospheric conditions are highly transient. The fluctuating liquid interface is translated within the focal plane of the laser beam, which produces the undesired outcome of fluctuations in the signal intensity measured by the mass spectrometer.<sup>23</sup> These fluctuations introduce the above-mentioned time correlation artifacts, so that time averaging is inappropriate, at least until the appropriate time correlation scales have been determined. Time-dependent mass spectroscopy has turned out to be much more challenging than initially expected.

Faced with such complications, one could opt for shorter runs with fewer data. Shorter acquisition times, however, make it harder to distinguish signal from noise and to disentangle peaks. There is no escape from the dilemma.

Fortunately, there is a way to obtain quantitative answers even when the data are sparse, multivariate, and/or complex. Encountering the same issues, fields such as particle physics and cosmology have increasingly applied the methods of Bayesian inference with success.<sup>24–26</sup> It therefore seems natural to us to apply similar methods to mass spectrometry data. Some Bayesian research has already been done with regard to separating joint mass and charge-dependent spectra into zero-charge mass distributions, and the related open-source software dubbed UniDec by the authors is available.<sup>27</sup>

In this paper, we investigate the application of Bayesian methods to time-of-flight mass spectrometry data. Our two main objectives are to test the use of Bayesian deconvolution to improve mass resolution of peaks and to evaluate the robustness of the methodology. The improved resolution of a deconvolved signal permits more quantitative statements regarding the fragmentation pathway of a well-known thermometer ion during femtosecond ultraviolet desorption. Once the proposed deconvolution methods have been stress-tested in this simple environment, they may be extended to assist in the decoupling of shot-to-shot phenomena.

As a valuable candidate measurement to examine the potential applicability and value of developing this framework, we selected a fragmentation pathway study, which characterizes the transfer of internal energy occurring between the matrix and the analyte ions during the laser desorption. Although matrix-assisted laser desorption/ionization (MALDI) is routinely used,<sup>28</sup> the underlying ionization mechanisms are not adequately understood and therefore remain a topic of active investigation.<sup>29–31</sup> Furthermore, since the internal energy of an ion determines the potential fragmentation pathways, it is a useful property for characterizing the softness of a given ionization mechanism.<sup>32–37</sup> However, due to the particular nature of these measurements, it is not experimentally appropriate to increase the mass resolution by implementing delayed ion extraction or an ion reflectron, since either of these interventions would eliminate valuable quantitative information from the mass spectra.<sup>38–40</sup> For example, reflecting analyzers impede the assignment of the fragments produced within the ionization and acceleration regions, while also introducing undesirable mass discrimination effects.<sup>41–43</sup> We therefore decided to use, as a case study, the linear mass spectra that resulted from the collisional transfer of energy during the laser desorption process from the 2,5-dihydroxybenzoic acid (DHB) matrix to the benzyltriphenylphosphonium (BTP) thermometer ion.

The relevant experiments were performed on our in-house designed linear time-of-flight mass spectrometer with a mass resolution of approximately 200 in the mass range of interest. As this system has been discussed before,<sup>44</sup> we sketch the relevant information only briefly. The third harmonic ( $\lambda = 343$  nm,  $\tau = 190$  fs) output pulses from a regeneratively amplified Yb:KGW oscillator (Pharos SP1.5, Light Conversion, Vilnius, Lithuania) was used for sample irradiation. Desorption was performed in the transmission geometry, after which the positive ions were accelerated to the nominal kinetic energy of 5 keV in a static, two-stage extraction region, supplemented by a 10 keV postacceleration stage reaching the detector. Positive ions were detected with a dual-stage chevron microchannel plate detector (F9890, Hamamatsu, Bridgewater, USA). These ion signals were recorded by an 8-bit digitizer (DC211, Acqiris, Plan-les-Ouates, Switzerland) operating at a 1 ns sampling rate. For all the data presented, 100 single-shot spectra were averaged before performing the deconvolution. The final results were normalized relative to the 2,5-dihydroxybenzoic acid (DHB)<sup>45</sup> matrix parent ion.

## II. METHODOLOGY

In the [supplementary material](#), we derive in detail our Bayesian deconvolution methods for extracting the underlying signals from low-resolution, experimental time-of-flight data and compare them to several existing methods. Due to the lengthiness of the derivations, we here highlight only the salient points. Our approach is based on the Lucy-Richardson (Poisson-likelihood) deconvolution algorithm<sup>46,47</sup> which we supplement by the inclusion of various Bayesian prior distributions using the split-gradient method.<sup>48,49</sup>

The mass spectrum data are represented as a set of discrete counts  $n_b$ , one for each  $m/z$  interval, channel, or bin  $b$ , where the joint intervals of bins  $b = 1, 2, \dots, B$  cover the entire  $m/z$  interval. The individual events counted in each bin during the acquisition follow a

Poisson distribution<sup>50</sup>  $p(n_b | \lambda_b) = e^{-\lambda_b} \lambda_b^{n_b} / n_b!$ , where the parameter  $\lambda_b > 0$  represents the signal intensity; it is proportional to the acquisition time period. The vector  $\mathbf{n} = \{n_b\}_{b=1}^B$  therefore represents the time-aggregated data of the entire spectrum. Assuming that counts  $n_b$  are mutually independent, the joint probability of all counts given the parameters  $\boldsymbol{\lambda} = \{\lambda_b\}_{b=1}^B$ , also termed the likelihood, is given by the following:

$$p(\mathbf{n} | \boldsymbol{\lambda}) = \prod_{b=1}^B p(n_b | \lambda_b). \quad (1)$$

The signal intensity  $\lambda_b$  is, however, the result of spectral broadening, i.e., the sum of contributions from narrower underlying peaks which could not be distinguished during measurement and may also partly fall into other nearby bins. Mathematically speaking,  $\lambda_b$  is the result of a convolution. The task at hand is to deconvolve the spectrum, i.e., to reverse that convolution so as to separate, as well as possible, the low-resolution data into high-resolution peaks, using, where possible, other pertinent information such as known detector responses. The objective is to find a set of parameters,  $\mathbf{s}^* = \{s_b^*\}_{b=1}^B$ , which represent the optimal amplitude approximations for each bin  $b$  while interpreting any small  $s_b^*$  as background noise.

Convolution and deconvolution are modeled as linear processes. Define  $s_c$  to be the true peak amplitude in bin  $c$  and let  $\mathbb{A}$  be the matrix whose components  $\mathbb{A}_{bc}$  constitute the peak broadening contribution which  $s_c$  makes to the data in nearby bins  $b$ . The Poisson parameter is then given by

$$\lambda_b = \sum_c \mathbb{A}_{bc} s_c \quad (2)$$

or  $\boldsymbol{\lambda} = \mathbb{A}\mathbf{s}$  in vector-matrix notation. Following the approximations set out in the [supplementary material](#), the likelihood is reduced to a variant of the Kullback-Leibler divergence called the  $I$ -divergence, which is the consistent measure for images and non-negative data in general.<sup>51</sup> By enforcing normalization, the negative logarithm of the likelihood  $L$  can be expressed as the so-called  $I$ -divergence,  $L[\mathbf{s}] = -\log p(\mathbf{n} | \boldsymbol{\lambda} = \mathbb{A}\mathbf{s}) = I[\mathbf{n} | \mathbb{A}\mathbf{s}]$ . The  $I$ -divergence, also called the relative entropy, replaces the metric distance which appears in the least-squares method. The desired deconvolved parameters  $\mathbf{s}^*$  are obtained from its minimization.

Our derivation in the [supplementary material](#) initially closely follows the Lucy-Richardson algorithm. However, the latter does not make use of priors, and priors are essential within a Bayesian framework in general and specifically for  $\mathbf{s}$ . Correspondingly, we are tasked to maximize not only the likelihood but the joint probability or equivalently to minimize the negative joint probability  $J[\mathbf{s}] = -\log p(\mathbf{n}, \mathbf{s})$ , the sum of the negative log likelihood  $L$ , and log prior  $P[\mathbf{s}] = -\log p(\mathbf{s})$ . As explained more fully below, a regularization parameter  $\beta$  is introduced which is the strength of the likelihood relative to the prior and is conventionally a coefficient of the prior.<sup>52</sup> The minimization is hence performed on

$$J[\mathbf{s}] = L[\mathbf{s}] + \beta P[\mathbf{s}]. \quad (3)$$

The system can be iteratively solved using the split-gradient method (SGM).<sup>48,49</sup> Due to the influence of the prior, the SGM in a multiplicative form does not always converge, in which case, the regularization parameter  $\beta$  has to be reduced and thereby the algorithm

comes to resemble the Lucy-Richardson algorithm<sup>53</sup> more closely which has no prior at all.

We now turn to the construction of a range of suitable priors reflecting possible different information available to the practitioner. A computationally simple family of priors is provided by the Gaussian distribution<sup>52</sup> with different choices for the so-called Toeplitz matrix  $\mathbb{B}_j$  and a scale parameter  $\Lambda$ . The Toeplitz matrix is the mathematical representation of a fixed function applied to each bin position. For example, the first discrete difference is the difference in the signal between two neighboring bins and can be represented by a Toeplitz matrix. The scale parameter  $\Lambda$  and the regularization parameter  $\beta$  have the same effect on the algorithm and are thus not independent. However, for pedagogy, we keep these parameters separate, as it is then easier to design a clear regularization strategy for  $\beta$ . We will set  $\Lambda$  to its expected value on each iteration such that it acts as a natural normalization of the prior. We, therefore, implemented the family of priors

$$p(\mathbf{s} | \Lambda, \mathbb{B}_j) = \left( \frac{\Lambda}{2\pi} \right)^{B/2} e^{-\Lambda \mathbf{s}^T \mathbb{B}_j \mathbf{s} / 2} \quad (4)$$

for different matrices  $\mathbb{B}_j$  reflecting the *a priori* information as follows: we may either have knowledge that the prior distribution depends on the signal itself, or we may know that it depends on a discrete difference between signals or even on a higher-order difference. The choice of  $\mathbb{B}_0 = 0$  then reflects the choice of having no prior at all, while dependence on the signal itself would motivate the usage of  $\mathbb{B}_1 = \mathbb{I}$  or, if a constant function is preferred, the first differences of the signal  $\mathbb{B}_2$ . Choices to use higher-order signal differences are reflected in the corresponding second-order difference matrix  $\mathbb{B}_3$  or the third-order  $\mathbb{B}_4$  difference.

In addition to generalizing the Lucy-Richardson algorithm, which is based on a Poisson-distributed likelihood, we have also implemented an equivalent methodology for Gaussian likelihoods (ISRA);<sup>54</sup> see the last section of the [supplementary material](#). The last part of our methodology, namely, our stopping criterion monitoring the mean of the residual distribution, and the role of the regularization parameter  $\beta$  are discussed in the numerical implementation section.

### III. NUMERICAL IMPLEMENTATION AND CHALLENGES

We here briefly discuss our algorithm, its usage, and the arising challenges. The pseudocode for the algorithm and its variables are discussed in the [supplementary material](#). Its implementation requires the user to make decisions regarding the perceived shape of the point spread function (PSF), the prior distribution, the usage of either Poisson (Lucy-Richardson) or Gaussian (ISRA) statistics, and the stopping criterion. These choices depend primarily on the spectra considered. The effectiveness of these options is judged by the speed of convergence and the quality of the recovered spectra. Here, we discuss the implementation of the algorithm and explain the choices to be made. In addition, the algorithm requires choices for the following:

1. The algorithm is only semiconvergent. In other words, after a certain number of iterations, the algorithm will distort the original signal and the model deteriorates. The most common solution is to implement a *stopping criterion*.



2. A useful *prior distribution* is required. In general, the prior which achieves convergence in the least number of iterations is the most optimal in the absence of other information.
3. A *regularization strategy* for the parameter  $\beta$  must be implemented to ensure convergence of the SGM.

The primary inputs into the algorithm consist of the initial values for the signal in the **Signal** vector and the data in the **Data** vector; the two vectors must be equal in length. Since the initial values of the **Signal** vector do not significantly affect the performance, we initialized them with ones. Other inputs are the vectors **Prior** and **Peak** which represent convolution kernels, i.e., representations of the Toeplitz matrices  $\mathbb{A}$  and  $\mathbb{B}$  to be convolved with the **Signal** vector. **Signal** is iteratively updated until the convergence criteria are satisfied, while the vectors **Data**, **Peak**, and **Prior** remain unchanged. The update depends on the convolution of the current **Signal** with the **Peak** and **Prior** vectors. Considering that in this paper the peak shapes of the present data are unimodal, and the results do not depend strongly on the peak shape, the shape vector **Peak** was set to a symmetric second-order polynomial.

While the relative entropy  $S$  between the model and data could be used as a stopping criterion, this proved challenging to generalize since its behavior depends strongly on the prior. Fortunately, the distribution of residuals as discussed in the deconvolution literature asserts that, for a meaningful reconstruction, the residuals should follow a Gaussian distribution centered around zero.<sup>26</sup> For example, the distribution of residuals is shown in Fig. 1(a) for the second-differences prior, while Fig. S1 depicts the reconstructed mass spectrum. Furthermore, Fig. S2 compares the deconvolved mass spectra obtained by the Lucy-Richardson and the second-difference algorithms. The irregular convergence behavior seen in Fig. 1 for the second-differences algorithm can be explained as the competition between the prior (which prefers flat structures) and the likelihood (which prefers spiked structures) causing very strong local domains in the parameter space.

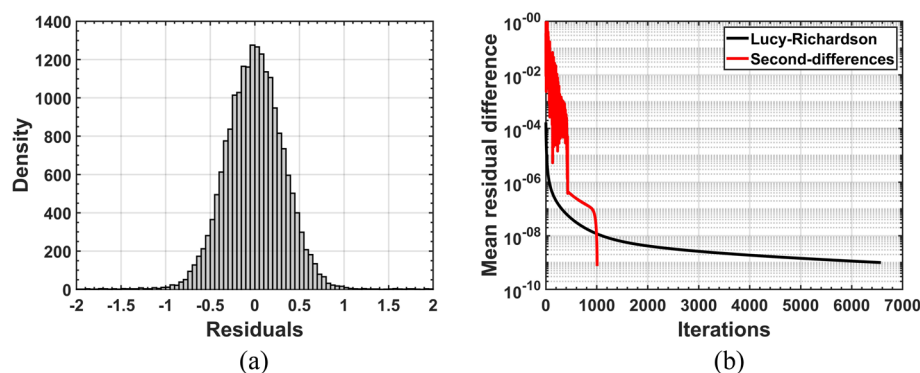
Since this distribution satisfies the above-mentioned requirements, we decided to utilize residuals and their behavior to monitor the state of the deconvolution process. Our stopping criterion requires that the *difference in the mean of the residual distribution*  $\epsilon'(i)$  and the *difference in the entropy*  $S'(i)$  must decrease below their predefined tolerance thresholds ( $\epsilon_{\text{TOL}}$  and  $S_{\text{TOL}}$ ) before reaching the maximum allowed number of iterations ( $I_{\text{MAX}} = 10^4$ ). In addition, a constant-sign criterion is implemented to counter

oscillations of the *second difference of the entropy*  $S''$  around zero. After an empirically determined number of oscillations ( $S_{\text{MAX}} = 4$ ), the strength of the prior is reduced, which accordingly eliminates such oscillatory behavior. Seeing that the magnitude of the *mean of the residuals*  $\epsilon(i)$  is determined by each dataset and prior, it cannot be employed as a universal stopping criterion; instead, the change  $\epsilon'(i)$  should be employed. The respective convergence behavior of the standard Lucy-Richardson algorithm ( $\mathbb{B}_0$ ) and the second-differences prior ( $\mathbb{B}_3$ ) are compared in Fig. 1. We confirmed that these trends agree qualitatively with the behavior of the entropy convergence plots.

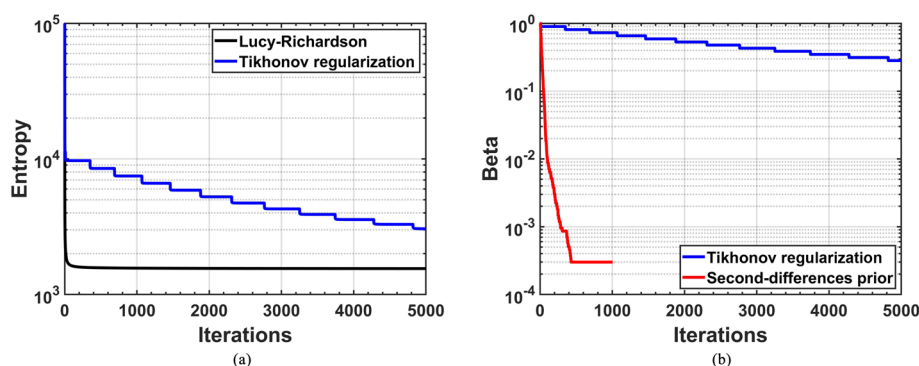
The difference in entropy  $S'(i)$  decrease on approaching local and global minima. When the difference falls below a predefined tolerance  $S_{\text{TOL}}$ , the strength of the prior is reduced by scaling the regularization parameter  $\beta$  by an amount of  $\beta_{\Delta} < 1$ , such that  $\beta' = \beta\beta_{\Delta}$  for the next iteration. The attenuation of the prior is used to weaken the influence of the local minima on the solution and is akin to the boosting mechanism introduced by Miroslav. Boosting is an *ad hoc* sharpening of the solution to move the solution out of the current local minima, while the prior attenuation is a weakening of the local minima, but with the same goal.<sup>55</sup>

Figure 2(b) indicates that the prior guides the algorithm during the initial phase of the deconvolution, while the influence of  $\beta$  is rather significant, but is increasingly suppressed by the boosting mechanism. We further note that without such a boosting mechanism the Tikhonov regularization prior ( $\mathbb{B}_1$ ) would have been permanently stranded within the local minimum<sup>56</sup> as indicated by the stepping behavior of the entropy in Fig. 2(a). Most importantly, this mechanism ensures that all the algorithms converge to a similar entropy, thereby permitting a quantitative comparison of results as a function of different priors for a given dataset. This consistency provides additional reassurance that the selected stopping criterion is indeed appropriate.

UniDec<sup>27</sup> is a specific implementation of the Lucy-Richardson algorithm to deconvolve the multiplicities of mass-to-charge ratios to find the underlying zero-charge mass distributions. By contrast, our approach is to augment both the Lucy-Richardson and the ISRA algorithms with prior information and regularizers such that the deconvolution of the mass-to-charge data with the instrument response function will improve the overall signal-to-noise ratio. It should be emphasized that the incorporation of a prior is not beneficial by default. Compared to the standard Lucy-Richardson algorithm, only the second-differences prior improved the rate of



**FIG. 1.** A representative residual error distribution (a) for the second-differences prior. The difference in the mean of residuals  $\epsilon'(i)$  is also shown (b) for this prior and standard Lucy-Richardson. The prior reaches the same mean residual difference as Lucy-Richardson, but in considerably fewer iterations while preserving the quality of the reconstruction. The algorithms terminated upon reaching the difference in the mean of the residual distribution tolerance  $\epsilon_{\text{TOL}}$  (additional iterations result in overfitting).



**FIG. 2.** Comparison of convergence of the entropy for the standard Lucy-Richardson and Tikhonov prior algorithms (a). Furthermore, the evolution of the regularization parameter  $\beta$  of the second-differences and Tikhonov priors is also shown (b). The influence exerted by these priors in the early phases is rapidly suppressed by the boost mechanism. For visual clarity, only the first 5000 of a much larger number of iterations are shown. All the algorithms terminated upon reaching the  $\varepsilon_{\text{TOT}}$  tolerance.

convergence from 952 to 223 iterations. The Tikhonov regularization prior ( $\mathbb{B}_1$ ) yielded the best spectra, avoiding overfitting of spurious small peaks and correctly identifying most of the physically meaningful peaks, but it did so at the expense of a markedly longer convergence time than the Lucy-Richardson algorithm (2024 vs 952 iterations). Both the first-difference ( $\mathbb{B}_2$ ) and third-differences ( $\mathbb{B}_4$ ) priors resulted in ill-defined peak shapes, which was not unexpected since these priors discourage sharp features (i.e., dependence on signal differences) within the solution.

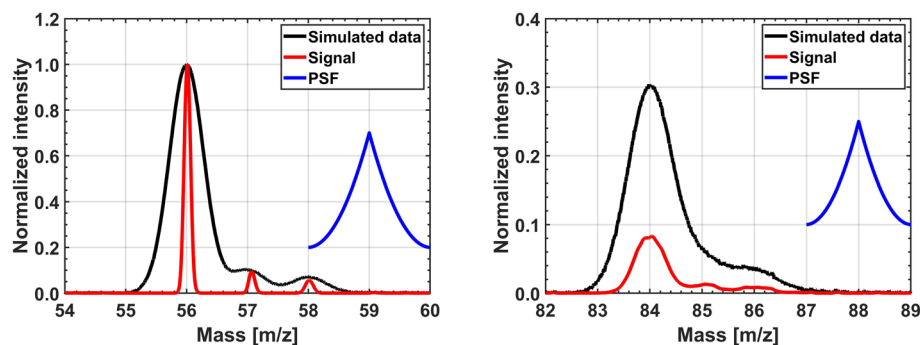
In order to better understand the performance of the algorithms, various simulated mass spectra were characterized (see the [supplementary material](#) for more details). Our most appropriate observation pertains to the effect of baselines. In an ideal scenario, without a baseline, Lucy-Richardson outperforms the second-differences prior algorithm significantly (343 vs 1576 iterations, respectively). In the presence of a baseline, however, the performance of the Lucy-Richardson algorithm deteriorates dramatically compared to the second-differences prior (6574 vs 1007 iterations). Both algorithms suffer from the tendency to overfit the baseline with a copious number of small, artifact peaks. Nonetheless, the second-differences prior does so at a considerably reduced rate and terminates earlier due to our novel stopping criterion. Figure S3 illustrates these observations.

Furthermore, all of the discussed algorithms suffer the limitation of using a peak shape function (PSF) which is independent of the charge-to-mass ratio. It is known that peak broadening increases with  $m/z$ . Implementing a position-dependent PSF would require a more sophisticated algorithm which is beyond the scope of this investigation.

Given that a predefined width is selected for the chosen PSF, only peaks within the data which are of a comparable width will be successfully deconvolved. As soon as a peak within the data is wider than the PSF, the algorithm will attempt to use multiple kernels during the deconvolution. It is therefore advantageous to select the PSF which is most appropriate for the given dataset. A common approach is to identify the most relevant mass range which needs to be deconvolved, and then to determine the corresponding PSF which best matches the peak widths in that range.

Peak width and the stopping criterion are also linked. The stopping criterion is defined in terms of the reconstruction error which monitors the error distribution averaged over all peaks rather than individual ones. This implies that peaks which are wider than the PSF kernel will not be adequately deconvolved because the algorithm is terminated well before the wider peaks are deconvolved as further iterations would degrade the already deconvolved peaks.

Figure 3 demonstrates by example the effect of a fixed-width PSF on



**FIG. 3.** Simulated silicon cluster data ( $\text{Si}_2$  and  $\text{Si}_3$ ) are used to demonstrate the effect of a fixed-width PSF on the deconvolution process, where the peak of interest (56 m/z) is successfully deconvolved, while the peak with wider features (84 m/z) is not. The PSF (inset) was optimized for the peak of interest.

two spectra of silicon clusters: the peak of interest at 56  $m/z$  is appropriately deconvolved, while the peak with wider features at 84  $m/z$  is not.

Although ion counting was performed, it was not known which error distribution would be more appropriate for this data since only the mass average is considered in this analysis. For that reason, we tested both Poissonian and Gaussian error statistics in the form of the Lucy-Richardson and the ISRA algorithms, respectively. As noted by others, and as further shown in Fig. S4 in the [supplementary material](#), Lucy-Richardson converges significantly faster than ISRA (223 vs 2530 iterations) while also yielding superior deconvolution results; the ISRA algorithm is more susceptible to introducing artifactual peaks. Currently, it is not known whether this is related to the specific stopping criterion used by us or whether this conclusion applies universally to these algorithms; certainly, the residual distributions appear to be qualitatively similar. Even though the results of not using priors are shown, these characteristics (superior performance of Lucy-Richardson over ISRA) were observed for all of the priors investigated. Preliminary results indicate that this observation also applies to single-shot spectra.

#### IV. RESULTS AND DISCUSSION

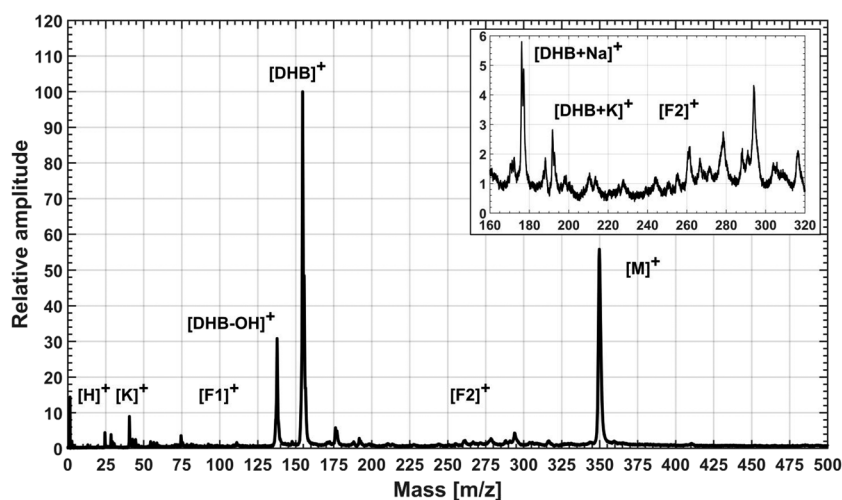
We have applied split-gradient deconvolution methods to laser desorption mass spectrometry time-of-flight data; the experimental setup has been described in Ref. 44. A representative example of the spectra obtained when investigating a BTP fragmentation pathway is shown in Fig. 4; both the parent peaks of the BTP thermometer ( $m/z$  353) and DHB matrix ions ( $m/z$  154) are pronounced. The BTP fragmentation signature is known to primarily consist of the benzyl ( $m/z$  91) and triphenylphosphine ( $m/z$  262) ions.<sup>37</sup> However, since femtosecond pulses are used, the fragmentation is reduced such that these peaks are barely visible.<sup>44</sup> The soft nature of desorption with ultrashort pulses has been discussed elsewhere.<sup>36,37,44</sup>

The mass resolution is significantly improved using an appropriate deconvolution method, as shown in Fig. 5, for the standard Lucy-Richardson algorithm and a modification thereof by applying the second-difference prior. The deconvolved signal facilitated

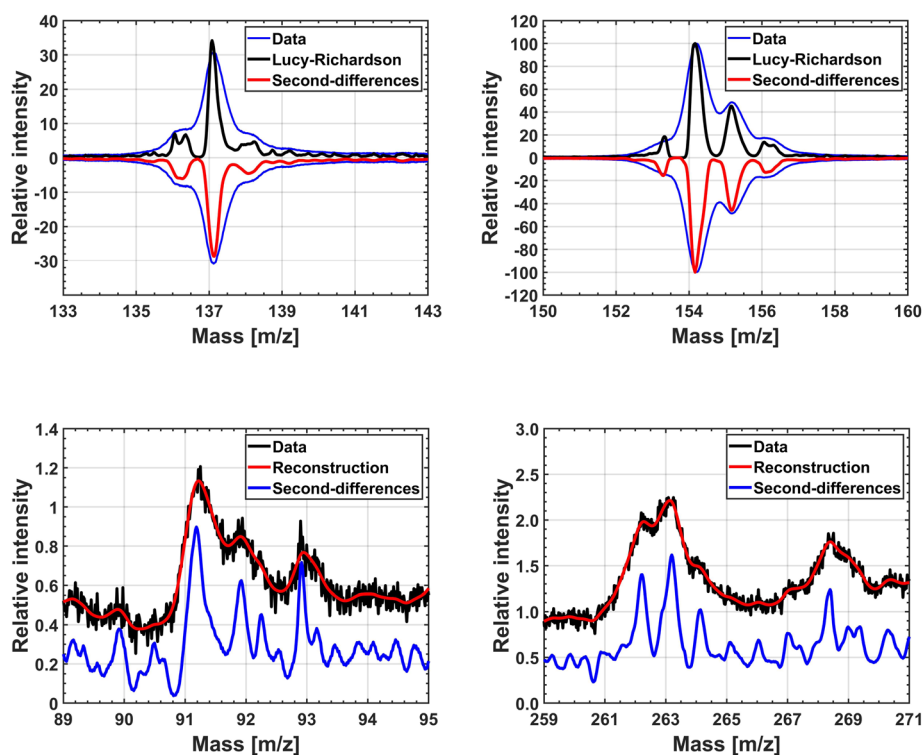
a more quantitative conclusion regarding the fragmentation pathway as it resulted in an improved resolution and, therefore, enabled more peak-to-peak ratios to be defined. Of the five priors investigated ( $\mathbb{B}_0$  to  $\mathbb{B}_4$ ), only the standard Lucy-Richardson algorithm ( $\mathbb{B}_0$ ) and a second-differences prior ( $\mathbb{B}_3$ ) are discussed in this section since these two methods produced the best results. Both of these methods deconvolved the data into the underlying signals while the peak amplitude ratios are evidently preserved. However, the second-differences prior performed better than Lucy-Richardson in deconvolving the underlying peak structures such as dehydroxylated DHB. This superior performance was observed for the majority of peaks, especially for those having relatively small amplitudes such as triphenylphosphine. For both of these methods, the recovered spectra (see Fig. 6) overlapped the raw data well, which signifies an appropriate deconvolution since the deconvolved signal is capable of recovering the data.

The Lucy-Richardson algorithm is semiconvergent after an initial deconvolution period, which suggests that further iterations will not substantially improve the likelihood while continuing to increase the sparsity of the solution. By adding the prior, the initial deconvolution is guided closer to an appropriate solution which accelerates the convergence and, therefore, also decreases the number of iterations that can introduce sparsity, thus preserving the peak amplitudes ratios. As an illustrative example, Fig. S5 (see [supplementary material](#)) shows two histograms comparing the respective preservation of the peak amplitude ratios, defined as the relative difference between the signal and data. Even though these two distributions have similar centroids, the distribution of the prior method is skewed toward the lower-end, thereby indicating the better relative peak amplitude preservation. Subsequent studies will investigate whether the isotopic distributions are equally well preserved.

Due to the minimal fragmentation of BTP, we will primarily discuss the DHB fragments. Our immediate goal was to enhance confidence in our methods for accurately quantifying ion fragmentation. However, given the low mass resolution ( $\sim 200$ ), it was difficult to succeed in doing this before the deconvolution. For example, examining the region following the DHB parent ion, the isotopic



**FIG. 4.** Representative spectrum after averaging several single-shot spectra. Although the parent peaks of the BTP thermometer ( $m/z$  353) and the DHB matrix ( $m/z$  154) ions stand out, the benzyl (F1,  $m/z$  91) and triphenylphosphine (F2,  $m/z$  262) ions are barely visible.



**FIG. 5.** Comparison spectra of DHB ( $m/z$  154, right) and its dehydroxylated fragment ( $m/z$  137, left) after deconvolution was performed using the standard Lucy-Richardson algorithm ( $\mathbb{B}_0$ ) and the Lucy-Richardson algorithm with the second-differences prior ( $\mathbb{B}_3$ ).

**FIG. 6.** Comparison spectra of the benzyl fragment ( $m/z$  = 91, fragment F1, left), as well the triphenylphosphine fragment ( $m/z$  = 261, fragment F2, right) ion after deconvolution was performed using the Lucy-Richardson algorithm with the second-differences prior.

distribution thereof can be approximated. Nevertheless, it is hard to make a realistic statement regarding the hydrogen loss peaks ( $m/z$  136 and 153) other than inferring their possible existence. Likewise, very few conclusions can be drawn regarding the dehydroxylated DHB fragment ( $m/z$  137) other than to determine its peak intensity ratio relative to the parent. At the minimum, the deconvolution appears to be successful in that each peak in Fig. 5 is separated by single atomic mass units. The hydrogen loss is supported by the  $H^+$  peak that is visible in Fig. 4, which suggests that a fast ejection of neutral hydrogen, followed by an ionization step, is probably the responsible pathway.<sup>57–60</sup> While similar pathways have been reported in previous studies when using ultrashort pulses, they could not be corroborated for these measurements before the deconvolution.

The base peak in Fig. 4 resembles the intact BTP cation ( $m/z$  = 353, parent  $M$ ), which is rationalized by the matrix-assisted desorption of the preformed cation from the solid state. It is noteworthy that the BTP fragmented only marginally, as only the peaks associated with the benzyl ( $m/z$  = 91, fragment F1) and triphenylphosphine ( $m/z$  = 261, fragment F2) radicals were observed. Unsurprisingly, these two fragment ions are characteristic of the intact cation, which is formed by unimolecular decomposition due to a homolytic cleavage.<sup>33,34</sup> Even though the relative intensities of both the benzyl ( $\sim 1\%$ ) and triphenylphosphine ( $\sim 2\%$ ) ions are very low, the data did suggest their presence. However, when comparing the deconvolved signal to the data, Fig. 6 shows that deconvolution enabled an enhanced peak identification because of this increase in the mass resolution ( $\sim 500$  at  $m/z$  154). It is important to recognize

that this improvement is comparable to that offered using delayed ion extraction.<sup>39,40</sup>

It was possible to recognize the benzyl fragment before deconvolution, but the enhanced mass resolution enabled its protonated species to also be identified. Mass separation suggests that the two other adjacent peaks ( $m/z$  = 92 and 93) are associated with one another, but not with the benzyl fragment. Similarly, only after the deconvolution, it was possible to interpret the mass signature of the triphenylphosphine fragment. It is imperative to highlight that the intensity ratio of the protonated species to that of the radical ion is larger than one. A study on the mass spectral behavior of alkyltriphenylphosphonium salts demonstrated that this ratio could be considered as an indicator of the cation internal energy distribution.<sup>32</sup> Later studies on BTP reiterated this observation.<sup>33</sup> In addition, the characteristic ion ( $m/z$  = 167) associated with thermal degradation of BTP was not observed.<sup>33</sup> Combined with the finding that no other fragments have been observed, these observations reaffirm the hypothesis that minimal energy was transferred from the DHB matrix to the BTP thermometer ions.<sup>37,44</sup>

## V. CONCLUSIONS

We have demonstrated the benefits and advantages of deconvolving low-resolution time-of-flight data by using the well-established Poisson-based Lucy-Richardson algorithm both with and without Bayesian priors. Different priors were applied so as to extract a more meaningful signal from the experimental data. For the data investigated in this analysis, it was shown that the



Gaussian prior based on the second-differences of the signal outperforms the standard Lucy-Richardson algorithm, as evidenced by faster convergence and the preservation of a larger fraction of the peak amplitudes ratios. A novel stopping criterion that monitors the difference in the mean of the residuals was included which, in combination with a boosting mechanism, ensures that the algorithm does not get stuck in a local minimum and that all methods reach the same result. The Gaussian-based Image Space Reconstruction algorithm was also studied, as it was not known originally whether Gaussian statistics might be more appropriate for the experimental test dataset. We can confirm that, as noted previously by others, Lucy-Richardson converges more rapidly and is also less prone to overfitting. For all investigated methods, the improved mass resolution of the deconvolved signal allowed a more accurate statement to be made regarding fragmentation of both the 2,5-dihydroxybenzoic acid (DHB) matrix, as well as the benzyltriphenylphosphonium (BTP) thermometer ion, following femtosecond ultraviolet laser desorption. Further studies will extend the framework introduced in this work to assist in the interpretation of MALDI shot-to-shot phenomena.

## SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for a more complete account of the underlying mathematical and computational issues summarized in the main text.

## ACKNOWLEDGMENTS

Cornelius L. Pieterse and Michiel B. de Kock have contributed equally. The authors would like to thank Spencer Thomas (National Centre of Excellence in Mass Spectrometry Imaging, National Physical Laboratory, Teddington, United Kingdom) for valuable comments and suggestions. This work was supported by the Max Planck Society and, in part, by the Excellence Cluster Universe of the Technical University of Munich and the National Research Foundation of South Africa.

## REFERENCES

1. Y.-Z. Liang, P. Xie, and K. Chan, *J. Chromatogr. B* **812**, 53 (2004).
2. J. M. Wiseman, D. R. Ifa, Y. Zhu, C. B. Kissinger, N. E. Manicke, P. T. Kissinger, and R. G. Cooks, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18120 (2008).
3. S.-E. Ong and M. Mann, *Nat. Chem. Biol.* **1**, 252 (2005).
4. B. F. Cravatt, G. M. Simon, and J. R. Yates III, *Nature* **450**, 991 (2007).
5. E. P. Diamandis, *Mol. Cell. Proteomics* **3**, 367 (2004).
6. E. P. Diamandis, *J. Natl. Cancer Inst.* **102**, 1462 (2010).
7. A. E. Clark, E. J. Kaleta, A. Arora, and D. M. Wolk, *Clin. Microbiol. Rev.* **26**, 547 (2013).
8. L. A. McDonnell, G. L. Corthals, S. M. Willems, A. van Remoortere, R. J. van Zeijl, and A. M. Deelder, *J. Proteomics* **73**, 1921 (2010).
9. L. S. Eberlin, A. L. Dill, A. J. Golby, K. L. Ligon, J. M. Wiseman, R. G. Cooks, and N. Y. Agar, *Angew. Chem., Int. Ed.* **49**, 5953 (2010).
10. O. Golf, N. Strittmatter, T. Karancsi, S. D. Pringle, A. V. Speller, A. Mroz, J. M. Kinross, N. Abbassi-Ghadi, E. A. Jones, and Z. Takats, *Anal. Chem.* **87**, 2527 (2015).
11. C. A. Smith, G. O'Maille, E. J. Want, C. Qin, S. A. Trauger, T. R. Brandon, D. E. Custodio, R. Abagyan, and G. Siuzdak, *Ther. Drug Monit.* **27**, 747 (2005).
12. D. S. Wishart, T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat, and E. Dong, *Nucleic Acids Res.* **41**, D801 (2012).
13. I. Fournier, C. Marinach, J. Tabet, and G. Bolbach, *J. Am. Soc. Mass Spectrom.* **14**, 893 (2003).
14. A. Wortmann, T. Pimenova, S. Alves, and R. Zenobi, *Analyst* **132**, 199 (2007).
15. P. Chaurand, K. E. Schriver, and R. M. Caprioli, *J. Mass Spectrom.* **42**, 476 (2007).
16. Y. J. Bae, K. M. Park, and M. S. Kim, *Anal. Chem.* **84**, 7107 (2012).
17. A. Ingendoha, M. Karasa, F. Hillenkamp, and U. Giessmann, *Int. J. Mass Spectrom. Ion Processes* **131**, 345 (1994).
18. K. Dreisewerd, M. Schürenberg, M. Karas, and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Processes* **141**, 127 (1995).
19. H. Qiao, V. Spicer, and W. Ens, *Rapid Commun. Mass Spectrom.* **22**, 2779 (2008); e-print [arXiv:NIHMS150003](#).
20. I. S. Gilmore, *J. Vac. Sci. Technol., A* **31**, 050819 (2013).
21. N. L. Anderson and N. G. Anderson, *Mol. Cell. Proteomics* **1**, 845 (2002).
22. K. O. Zhurov, A. N. Kozhinov, L. Fornelli, and Y. O. Tsybin, *Anal. Chem.* **86**, 3308 (2014).
23. Y. Lu, C. L. Pieterse, W. D. Robertson, and R. J. D. Miller, *Anal. Chem.* **90**, 4422 (2018).
24. J. Skilling, *Maximum Entropy and Bayesian Methods* (Springer Science & Business Media, Cambridge, England, 2013), Vol. 36, p. 1988.
25. R. Narayan and R. Nityananda, *Annu. Rev. Astron. Astrophys.* **24**, 127 (1986).
26. R. Puettner, T. Gosnell, and A. Yahil, *Annu. Rev. Astron. Astrophys.* **43**, 139 (2005).
27. M. T. Marty, A. J. Baldwin, E. G. Marklund, G. K. A. Hochberg, J. L. P. Benesch, and C. V. Robinson, *Anal. Chem.* **87**, 4370 (2015).
28. M. Karas and F. Hillenkamp, *Anal. Chem.* **60**, 2299 (1988).
29. K. Dreisewerd, *Chem. Rev.* **103**, 395 (2003).
30. Y. J. Bae and M. S. Kim, *Annu. Rev. Anal. Chem.* **8**, 41 (2015).
31. R. Knochenmuss, *Annu. Rev. Anal. Chem.* **9**, 365 (2016).
32. D. A. McCrery, D. A. Peake, and M. L. Gross, *Anal. Chem.* **57**, 1181 (1985).
33. J. Claereboudt, M. Claeys, H. Geise, R. Gijbels, and A. Vertes, *J. Am. Soc. Mass Spectrom.* **4**, 798 (1993).
34. J. Claereboudt, W. Baeten, H. Geise, and M. Claeys, *J. Mass Spectrom.* **28**, 71 (1993).
35. G. Luo, I. Marginean, and A. Vertes, *Anal. Chem.* **74**, 6185 (2002).
36. Y. Chen and A. Vertes, *J. Phys. Chem. A* **107**, 9754 (2003).
37. G. Luo, I. Marginean, L. Ye, and A. Vertes, *J. Phys. Chem. B* **112**, 6952 (2008).
38. B. A. Mamyrin, V. I. Karataev, D. V. Shmikk, and V. A. Zagulin, *Sov. Phys. - JETP* **37**, 45 (1973).
39. R. S. Brown and J. J. Lennon, *Anal. Chem.* **67**, 1998 (1995).
40. M. Vestal, P. Juhasz, and S. Martin, *Rapid Commun. Mass Spectrom.* **9**, 1044 (1995).
41. U. Boesl, R. Weinkauff, and E. Schlag, *Int. J. Mass Spectrom. Ion Processes* **112**, 121 (1992).
42. W. J. Jia, K. W. D. Ledingham, C. T. J. Scott, C. Kosmidis, and R. P. Singhal, *Rapid Commun. Mass Spectrom.* **9**, 761 (1995).
43. M. Vestal and P. Juhasz, *J. Am. Soc. Mass Spectrom.* **9**, 892 (1998).
44. C. L. Pieterse, F. Busse, F. Tellkamp, W. D. Robertson, and R. J. D. Miller, e-print [chemrxiv.6450380.v2](#).
45. K. Strupat, M. Karas, and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Processes* **111**, 89 (1991).
46. L. B. Lucy, *Astron. J.* **79**, 745 (1974).
47. W. H. Richardson, *J. Opt. Soc. Am.* **62**, 55 (1972).
48. H. Lantéri, M. Roche, O. Cuevas, and C. Aime, *Signal Process.* **81**, 945 (2001).
49. H. Lanteri, M. Roche, and C. Aime, *Inverse Probl.* **18**, 1397 (2002).
50. J. M. Bernardo and A. F. Smith, *Bayesian Theory*, Wiley Series in Probability and Statistics (Wiley, 1994).
51. I. Csizsar, *Ann. Stat.* **19**, 2032 (1991).
52. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1996), Vol. 2.

- <sup>53</sup>L. A. Shepp and Y. Vardi, *IEEE Trans. Med. Imaging* **1**, 113 (1982).
- <sup>54</sup>M. E. Daube-Witherspoon and G. Muehllehner, *IEEE Trans. Med. Imaging* **5**, 61 (1986).
- <sup>55</sup>M. Morháč, *Nucl. Instrum. Methods Phys. Res., Sect. A* **559**, 119 (2006).
- <sup>56</sup>A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems* (Springer Netherlands, Dordrecht, 1995), p. 257.
- <sup>57</sup>J. J. Yang, D. A. Gobeli, and M. A. El-Sayed, *J. Phys. Chem.* **89**, 3426 (1985).
- <sup>58</sup>R. Weinkauff, P. Aicher, G. Wesley, J. Grotemeyer, and E. Schlag, *J. Phys. Chem.* **98**, 8381 (1994).
- <sup>59</sup>D. Smith, K. Ledingham, H. Kilic, T. McCanny, W. Peng, R. Singhal, A. Langley, P. Taday, and C. Kosmidis, *J. Phys. Chem. A* **102**, 2519 (1998).
- <sup>60</sup>A. N. Markevitch, N. P. Moore, and R. J. Levis, *Chem. Phys.* **267**, 131 (2001).