# Rapid Deconvolution of Low-Resolution Time-of-Flight Data using Bayesian Inference

Cornelius L. Pieterse,[†,‖] Michiel B. de Kock,[‡] Wesley D. Robertson,[†]

Hans C. Eggers,[‡,¶] and R.J. Dwayne Miller[*,†,§]

†*Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany*

‡*Department of Physics, Stellenbosch University, Matieland, 7602, South Africa*

¶*National Institute for Theoretical Physics, Matieland, 7602, South Africa*

§*Departments of Chemistry and Physics, University of Toronto, 80 St. George Street, Toronto, Ontario, M5S 3H6, Canada*

‖*Present address: National Centre of Excellence in Mass Spectrometry Imaging, National Physical Laboratory, Hampton Road, Teddington, TW11 0LW, United Kingdom*

E-mail: 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 by various Bayesian prior distributions and show that a prior of second-differences of the signal outperforms the standard Lucy-Richardson algorithm by accelerating the rate of convergence by a factor of two and preserving the peak amplitude ratios of a larger fraction of the total peaks. A stopping criterion and boosting mechanism is implemented to ensure that these methods converge to the same final entropy and that local minima are avoided. Improvement by a factor of two in mass resolution of the signal allowed more accurate quantification of the spectra. The general method is demonstrated in this paper by the deconvolution of fragmentation pathway peaks of the benzyltriphenylphosphonium thermometer ion following femtosecond ultraviolet laser desorption.

# Introduction

Mass spectrometry (MS) is practiced in various settings, including commercial applications such as quality control[1] and pharmacokinetics,[2] as well as basic scientific investigations such as proteomics[3] and biological pathway analysis.[4] While 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[5–7] 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 mammalia tissue sections when comparing their respective mass spectrometric data.[8–10] This opens up the possibility of compiling a mass spectrometric database of biomarkers associated with recognized diseases for the identification of unhealthy tissue.[10–12] 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 data sets 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 within which it might be of interest to study the mass spectra as a function of depth or shot number.[13–16]

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 data sets 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.

It is of course also possible to obtain good results by brute force long acquisition times, but this may have other unintended consequences. As an example, more accurate measurements could be obtained by the use 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.[17–20] It is common knowledge that the concentration of the biomarkers of interest is usually orders of magnitude lower than that of the background, closely related mass species in the surrounding tissue, which frequently includes species which contribute to ion suppression and, therefore, hinder detection of these biomarkers.[3,21,22]

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 obscures 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.[23] 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 less data. Shorter acquisition times, however, make it harder to distinguish signal from noise and to disentangle peaks. There is no escape from the dilemma. There is, however, a way to obtain quantitative answers even when the data is 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.[24–26] It hence seems natural to us to apply similar methods to mass spectrometry images.

In this paper, we examine 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 this framework. The improved resolution of a deconvolved signal permits more quantitative statements regarding the fragmentation pathway of a well-known thermometer ion upon 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 such a Bayesian framework, a fragmentation pathway study was selected, which characterizes the internal energy transfer occurring between the matrix and the analyte ions during the laser desorption. Although matrix-assisted laser desorption/ionization (MALDI) is routinely used,[27] the underlying ionization mechanisms are not adequately understood and are therefore still recurrently investigated.[28–30] 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.[31–36] However, due to the specific nature of these measurements, it is not experimentally desirable to increase the mass resolution by means of implementing delayed ion extraction or an ion reflectron, since either of these interventions would eliminate valuable quantitative information from the mass spectra.[37–39] We, therefore, used as a case study the linear mass spectra produced by internal energy transferred during the desorption process 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,[40] 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 post-acceleration 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, Acquiris, Plan-les-Ouates, Switzerland) operating at an 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)[41] matrix parent ion.

# Methodology

In this section, we derive and discuss several Bayesian deconvolution methods for extracting the underlying signal from low-resolution time-of-flight data. Our approach is based on the well-established Lucy-Richardson deconvolution algorithm[42,43] which we supplemented with Bayesian prior distributions.

We show that a Gaussian prior based on the second-differences of the signal outperforms the standard Lucy-Richardson algorithm in terms of preserving the peak amplitude ratios for a larger fraction of the total number of peaks. To facilitate comparisons, a novel stopping criterion is introduced which monitors the difference in the mean of the residuals, which in combination with a boosting mechanism, ensures that the algorithm does not wind up in a local minimum and that all of the methods converge to the same result.

## Linear deconvolution

From a statistics viewpoint, one-dimensional mass spectrometric data is 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 $b = 1, 2, \ldots, B$ cover the entire $m/z$ interval. Barring other pertinent information, the individual events counted in any bin during the acquisition are considered *exchangeable*;[44] suggesting that the total bin counts $n_b$ follow a Poisson distribution

$$p(n_b \,|\, \lambda_b) = e^{-\lambda_b} \lambda_b^{n_b} / n_b! \qquad n_b = 0, 1, 2, \ldots, \infty \qquad b = 1, 2, \ldots, B, \tag{1}$$

where for each bin the parameter $\lambda_b > 0$, which is proportional to the acquisition time period, represents the true *signal* or expected value of counts in that bin. The vector $\boldsymbol{n} = \{n_b\}_{b=1}^{B}$, therefore, represents the time-averaged data of the experiment. 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

$$p(\boldsymbol{n} \,|\, \boldsymbol{\lambda}) = \prod_{b=1}^{B} p(n_b \,|\, \lambda_b). \tag{2}$$

By assumption, the vector $\boldsymbol{n}$ is the sum of counts originating from underlying, but spectrally broadened, narrow peaks: each $n_b$ is the *convolution* of these narrow peak counts. The task at hand is to reverse that convolution and to separate low-resolution data into high-resolution peaks using *deconvolution*, using where possible other pertinent information such as isotopic signatures or detector responses. The goal is to find a set of parameters $\boldsymbol{s}^* = \{s_b^*\}_{b=1}^{B}$ which represent the best amplitude approximations of a possible narrow unbroadened peak for each bin $b$, interpreting any small $s_b^*$ as background noise rather than a true peak. Research on deconvolution with Poisson likelihoods started during the mid-eighties after seminal papers by Shepp and Vardi[45] and by Geman and Geman[46]. A recent review of the literature appears in Bertero et al.[47] and reviews focussed on astronomy in Puetter et al.[26] and Starck et al.[48] Books dedicated to the subject are Hansen et al.[49], Jansson[50] and Young et al.[51]

Convolution and deconvolution are modelled as linear processes. Let $s_c$ be the true peak amplitude in bin $c$ and let $\mathbb{A}$ be the $B \times B$ square matrix whose components $\mathbb{A}_{bc}$ constitute the peak broadening contribution which $s_c$ makes to the data in nearby bins $b$. The matrix $\mathbb{A}$ is generally termed the point spread function (PSF). The Poisson parameter in bin $b$, in component and vector-matrix notation respectively, is then given by

$$\lambda_b = \sum_c \mathbb{A}_{bc} \boldsymbol{s}_c \qquad \text{or} \qquad \boldsymbol{\lambda} = \mathbb{A}\boldsymbol{s}. \tag{3}$$

Provided that the PSF depends only on the separation between the bins $b$ and $c$, the matrix $\mathbb{A}$ becomes a *Toeplitz matrix* whose components depend only on the difference between the row and column indices, $\mathbb{A}_{bc} = F_{b-c}$ where $F$ is some non-negative function, which we can write somewhat ambiguously as $\mathbb{A}_{b-c}$. The Poisson parameters can then be written as the convolution equation

$$\lambda_b = \sum_c F_{b-c}\, \boldsymbol{s}_c. \tag{4}$$

In addition, if the point spread function has finite support (i.e. the number of neighbouring bins $c$ contributing to $\lambda_b$ is finite), then the Toeplitz matrix $\mathbb{A}$ has a block-diagonal form.

Under these assumptions and limitations, the likelihood Eq. (2) can be rewritten as

$$p(\boldsymbol{n} \,|\, \mathbb{A}, \boldsymbol{s}) = \prod_b p(n_b \,|\, (\mathbb{A}\boldsymbol{s})_b) = \prod_b e^{-(\mathbb{A}\boldsymbol{s})_b}\, (\mathbb{A}\boldsymbol{s})_b^{n_b} \,/ n_b! \,. \tag{5}$$

Applying Stirling's approximation $\log n! \simeq n \log n - n$ to all counts $n_b$, the negative logarithm of the likelihood reduces to a variant of the Kullback-Leibler divergence,

$$L[\boldsymbol{s}] = -\log p(\boldsymbol{n} \,|\, \mathbb{A}, \boldsymbol{s}) \simeq \mathbf{1}^\mathsf{T} (\mathbb{A}\boldsymbol{s} - \boldsymbol{n}) + \boldsymbol{n}^\mathsf{T} \log \frac{\boldsymbol{n}}{(\mathbb{A}\boldsymbol{s})}, \tag{6}$$

where $\mathbf{1}^\mathsf{T}$ is a row vector of ones and for the notational simplicity we write $\boldsymbol{n}^\mathsf{T} \log \boldsymbol{n}/(\mathbb{A}\boldsymbol{s}) \equiv \sum_b n_b \log[n_b/(\mathbb{A}\boldsymbol{s})_b]$, i.e. the division and the logarithm are taken pointwise. This variant of the Kullback-Leibler divergence is called the I-divergence[52] which is the consistent measure for images and data which are non-negative. By enforcing the normalization of the matrix $\mathbf{1}^\mathsf{T}\mathbb{A} = \mathbf{1}^\mathsf{T}$, as appropriate for convolutions, this expression simplifies to

$$L[\boldsymbol{s}] = I[\boldsymbol{n} \,|\, \mathbb{A}\boldsymbol{s}] = \mathbf{1}^\mathsf{T} (\boldsymbol{s} - \boldsymbol{n}) + \boldsymbol{n}^\mathsf{T} \log \frac{\boldsymbol{n}}{\mathbb{A}\boldsymbol{s}}. \tag{7}$$

The I-divergence, or also called relative entropy, $I[\boldsymbol{n} \,|\, \mathbb{A}\boldsymbol{s}]$, replaces the metric distance that appears in the usual least-squares method and can be considered as the data fidelity term. It is convex, non-negative and coercive on the non-negative orthant (the higher-dimensional generalization of the octant), implying that a minimum exists which is global and unique. The gradient and Hessian of the I-divergence are given by

$$\nabla I[\boldsymbol{n} \,|\, \mathbb{A}\boldsymbol{s}] = \mathbb{A}^\mathsf{T} \left( \mathbf{1} - \frac{\boldsymbol{n}}{\mathbb{A}\boldsymbol{s}} \right), \qquad \nabla^2 I[\boldsymbol{n} \,|\, \mathbb{A}\boldsymbol{s}] = \mathbb{A}^\mathsf{T} \mathrm{diag} \left( \frac{\boldsymbol{n}}{(\mathbb{A}\boldsymbol{s})^2} \right) \mathbb{A}. \tag{8}$$

## Lucy-Richardson and Poisson algorithms

To solve the system of equations (8), the minimizer $\boldsymbol{s}^*$ of the I-divergence must be determined. The linear terms in the divergence imply that the solution must obey the constraint

$$\sum_b s_b^* = \sum_b n_b. \tag{9}$$

Moreover, the logarithm necessarily also requires that $s_b^* > 0$ for all $b$. As the convolution is a linear operation, perfect reconstruction of the data would, in general, require both negative and positive parameter values, and this positivity condition, therefore, complicates matters considerably. In effect, it forces minimizers of the I-divergence to be sparse, i.e. the solution $\boldsymbol{s}^*$ must lie near to the boundary of the non-negative orthant. This is called the *checkerboard effect*[53] or *night-sky reconstruction.*[54] If the underlying signal contains extended objects (i.e. the object spans more than one bin), this will conflict with the sparsity precondition and the algorithm should be stopped as soon as an appropriate solution is found. The algorithm is therefore semi-convergent. The main challenge is that the likelihood does not contain all the relevant information on what constitutes an image, and running it longer will only generate less plausible configurations.

Up until this point, the algorithm has followed the generally assumed superiority of Lucy-Richardson closely. To integrate it into a Bayesian framework, an appropriate prior for $\boldsymbol{s}$ must be specified. Instead of maximizing just the likelihood, the task becomes the maximization of the joint probability, which is the product of likelihood times source prior $p(\boldsymbol{n}, \boldsymbol{s}) = p(\boldsymbol{n} \mid \mathbb{A}, \boldsymbol{s})\, p(\boldsymbol{s})$, or equivalently minimization of the sum of the negative log likelihood and log prior $P[\boldsymbol{s}] = -\log p(\boldsymbol{s})$,

$$J[\boldsymbol{s}] = L[\boldsymbol{s}] + \beta P[\boldsymbol{s}], \tag{10}$$

where we have introduced a regularization parameter $\beta$ which mediates the strength of the likelihood relative to the prior. Conventionally, the regularization parameter is placed with the prior instead of the likelihood.[55] To solve the system iteratively, we apply the gradient descent method in a general form,

$$\boldsymbol{s}_{j+1}^* = \boldsymbol{s}_j^* - \alpha f[\boldsymbol{s}_j^*] \nabla J[\boldsymbol{s}_j^*], \tag{11}$$

where $\alpha$ is a relaxation factor, $f[\boldsymbol{s}_j^*]$ is the Lagrange function that ensures the constraint and $\nabla J[\boldsymbol{s}_j^*]$ is the gradient. If $\alpha$ is chosen appropriately and $f[\boldsymbol{s}_j^*]$ is a positive function within the domain of $\boldsymbol{s}^*$, each iteration will lower the I-divergence, and the process will converge. To enforce the positivity constraint, we set $f[\boldsymbol{s}^*] = \boldsymbol{s}^*$ such that a system that starts from a

positive solution will stay positive. It corresponds to performing the transform $\boldsymbol{s} = \exp[\boldsymbol{t}]$, seeking a minimum in $\boldsymbol{t}$ and then transforming back to $\boldsymbol{s}$. Inserting Eq. (8) gives

$$\boldsymbol{s}^*_{j+1} = \boldsymbol{s}^*_j - \alpha \boldsymbol{s}^*_j \left\{ \mathbf{1} - \mathbb{A}^\mathsf{T} \left( \frac{\boldsymbol{n}}{\mathbb{A}\boldsymbol{s}^*_j} \right) + \beta \nabla P[\boldsymbol{s}^*_j] \right\}. \tag{12}$$

Setting $\alpha = 1$ and without a prior, we have the multiplicative form of the algorithm, which is the standard Lucy-Richardson algorithm,

$$\boldsymbol{s}^*_{j+1} = \boldsymbol{s}^*_j \mathbb{A}^\mathsf{T} \left( \frac{\boldsymbol{n}}{\mathbb{A}\boldsymbol{s}^*_j} \right). \tag{13}$$

This multiplicative form explicitly enforces the positivity constraint and therefore reduces the computational requirements of the algorithm significantly. Our own numerical investigation has indicated this form has significant advantages over its competitors. The Lucy-Richardson algorithm was introduced independently by Lucy[42] and Richardson.[43] It was rederived by Sheppi and Verdi[45] as an example of the Expectation-Maximization(EM) algorithm,[56] which is itself a specific case of the Majorization-Minization approach. To include the prior, we need to split the gradient into positive and negative parts,

$$\nabla P[\boldsymbol{s}^*] = \boldsymbol{u}^* - \boldsymbol{v}^*, \tag{14}$$

where $\boldsymbol{u}^* \geq 0$ and $\boldsymbol{v}^* \geq 0$ for all $b$. Rewriting the derivative equation as

$$\boldsymbol{s}^* \left(1 + \beta \boldsymbol{u}^*\right) = \boldsymbol{s}^* \mathbb{A}^\mathsf{T} \left( \frac{\boldsymbol{n}}{\mathbb{A}\boldsymbol{s}^*} \right) + \beta \boldsymbol{v}^* \tag{15}$$

which can then be solved iteratively with

$$\boldsymbol{s}^*_{j+1} = \boldsymbol{s}^*_j \left\{ \mathbb{A}^\mathsf{T} \left( \frac{\boldsymbol{n}}{\mathbb{A}\boldsymbol{s}^*_j} \right) + \beta \boldsymbol{v}^*_j \right\} / \left(1 + \beta \boldsymbol{u}^*_j\right). \tag{16}$$

This algorithm is called the split-gradient method (SGM).[57,58] The SGM in the multiplicative form is not always convergent due to the influence of the prior. Should this happen, then the regularization parameter $\beta$ needs to be reduced such that the algorithm is more aligned with the Lucy-Richardson algorithm.[45] In summary, the following issues need to be addressed:

1. A stopping criterion is required for the semi-convergence of the algorithm.

2. A useful prior distribution corresponding to the type of solutions we prefer is required.

3. The regularization parameter $\beta$ must be adjusted so that the SGM is convergent.

## Priors for the SGM

A computational simple prior distribution to consider is a Gaussian distribution[55] with the choice of an appropriate Toeplitz matrix $\mathbb{B}_j$ and a scale parameter $\Lambda$,

$$p(\boldsymbol{s} \,|\, \Lambda, \mathbb{B}_j) = \left(\frac{\Lambda}{2\pi}\right)^{B/2} e^{-\Lambda \boldsymbol{s}^{\mathsf{T}} \mathbb{B}_j^{\mathsf{T}} \mathbb{B}_j \boldsymbol{s}/2}. \tag{17}$$

The possible matrices $\mathbb{B}_j, j = 0, 1, 2, 3, 4$ reflect the underlying generic information as follows. We may have generic knowledge that the prior distribution depends either on the signal itself, or we may know that it depends on a discrete difference between signals, or even on higher-order differences. The choice of $\mathbb{B}_0 = 0$ then reflects the desire to have 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

$$\mathbb{B}_2 = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & & & & \vdots \\ 0 & \dots & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}. \tag{18}$$

Choices to use higher-order signal differences are reflected in the corresponding second-order difference matrix

$$\mathbb{B}_3 = \begin{bmatrix} -1 & 2 & -1 & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & & & & \vdots \\ 0 & \dots & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & \dots & 0 & 0 & 0 & -1 & 2 & -1 \end{bmatrix}, \tag{19}$$

or alternatively, even a third-order difference,

$$\mathbb{B}_4 = \begin{bmatrix} 1 & -4 & 6 & -4 & 1 & 0 & \dots & 0 \\ 0 & 1 & -4 & 6 & 4 & 1 & \dots & 0 \\ \vdots & & & \ddots & & & & \vdots \\ 0 & \dots & 1 & -4 & 6 & -4 & 1 & 0 \\ 0 & \dots & 0 & 1 & -4 & 6 & -4 & 1 \end{bmatrix}. \tag{20}$$

As indicated earlier, all these matrices are Toeplitz Matrices, which conveniently represent the convolution operation. As the scale parameter $\Lambda$ appears only in the prior, it need not be a variable but can be replaced by its expectation value,

$$\Lambda^* = \frac{B}{1 + s^* \mathbb{B}_j^\intercal \mathbb{B}_j \boldsymbol{s}^*}, \tag{21}$$

where $B$ is the total number of bins or elements. The $\Lambda^*$ parameter can be computed on each iteration and normalizes the prior contribution. The stopping criterion, which monitors the difference in the mean of the residuals, and regularization parameter $\beta$ are to be discussed below. The residuals are defined as the difference between the reconstructed data using our minimizers $\mathbb{A}\boldsymbol{s}^*$ and the time-of-flight experimental data.

## Gaussian algorithm

In order to investigate the effect of the underlying error distributions, we also implemented an alternative approach by simply replacing the Poisson likelihood with a Gaussian distribution. Following the same steps as in the Poisson case, we obtain an iteration prescription

$$\boldsymbol{s}_{j+1}^* = \boldsymbol{s}_j^* \left\{ \left( \frac{\mathbb{A}^\intercal \boldsymbol{n}}{\mathbb{A}^\intercal \mathbb{A} \boldsymbol{s}_j^*} \right) + \beta \boldsymbol{v}_j^* \right\} / \left( 1 + \beta \boldsymbol{u}_j^* \right). \tag{22}$$

Without the prior, the algorithm is called the Image Space Reconstruction Algorithm (ISRA) or Muller Algorithm, which was originally proposed by Daube-Witherspoon and Muehllehner[59] and later analyzed by De Pierro[60] and Titterington.[61]

# Results and discussion

## Experimental test case

We have applied split-gradient deconvolution methods to laser desorption mass spectrometry time-of-flight data; the experimental setup has been described in Ref. 40. A representative example of the spectra obtained when investigating a BTP fragmentation pathway is shown in Figure 1; 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.[36] However, since femtosecond pulses are used, the fragmentation is reduced such that these peaks are barely visible.[40] The soft nature of desorption with ultrashort pulses has been discussed elsewhere.[35,36,40]

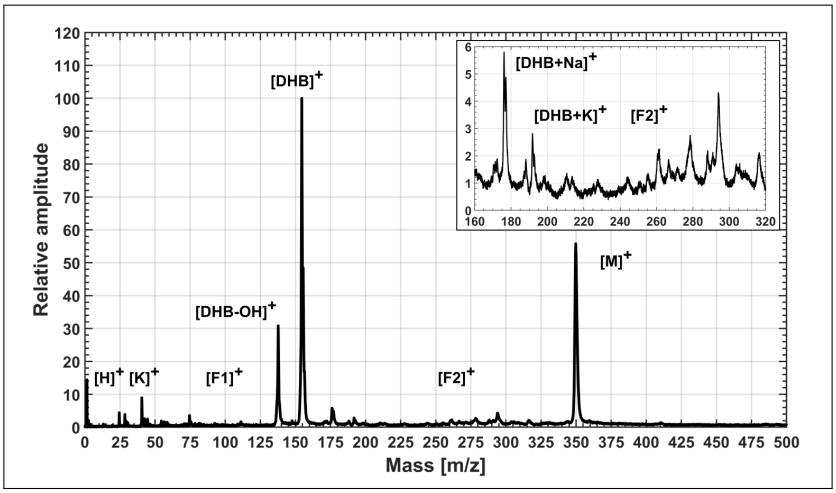



Figure 1: Representative spectrum after averaging several single-shot spectra. Although the parent peaks of the BTP thermometer ($m/z$ 353) and the DHB matrix ($m/z$ 154) ions stand out, the benzyl (F1, m/z 91) and triphenylphosphine (F2, m/z 262) ions are barely visible.

The mass resolution is significantly improved using an appropriate deconvolution method, as shown in Figure 2 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 these methods, the recovered spectra (not shown here) overlapped the data well, which signifies an appropriate deconvolution since the deconvolved signal is capable of recovering the data.

The Lucy-Richardson algorithm is semi-convergent 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, Figure S1 (see supporting information) 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 towards the lower-end, thereby indicating the better relative peak amplitude preservation. Subsequent studies will investigate whether the isotopic distributions are equally well preserved.

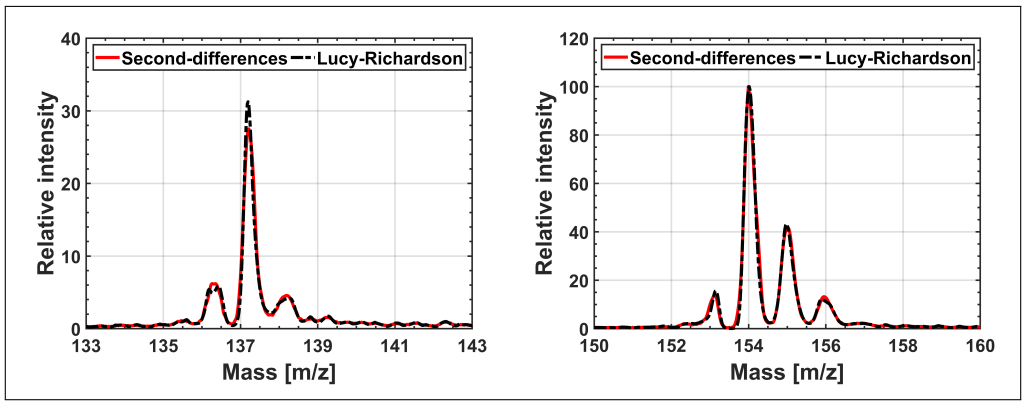



Figure 2: Comparison spectra of DHB (m/z 154, right) and its dehydroxylated fragment (m/z 137, left) after deconvolution was performed using the standard Lucy-Richardson algorithm ($\mathbb{B}_0$) and the Lucy-Richardson algorithm with the second-differences prior ($\mathbb{B}_3$).

Our immediate goal was to enhance confidence in our methods for accurately quantifying ion fragmentation. However, given the low mass resolution (∼200), it was difficult to succeed in doing this before deconvolution. As an example, examining the region following the DHB parent ion, the isotopic 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 of the peaks in Figure 2 are separated by single atomic mass units. The hydrogen loss is supported by the $H^+$ peak which is visible in Figure 1, which suggests that a fast ejection of neutral hydrogen, followed by an ionisation step, is probably the responsible pathway.[62–65] While similar pathways have been reported in previous studies when using ultrashort pulses,

they could not be corroborated for these measurements before deconvolution was performed. The benzyl fragment is still not clearly visible, suggesting that this fragment is suppressed, while it is at this time possible to identify the triphenylphosphine fragment by using its mass signature.[32,33] When comparing the deconvolved signal to the data, Figure 3 illustrates that the deconvolution facilitated an enhanced peak identification process because of an increase in the mass resolution (∼500 at m/z 154). It is important to recognize that this improvement is comparable in magnitude to that offered when using delayed ion extraction.[38,39]

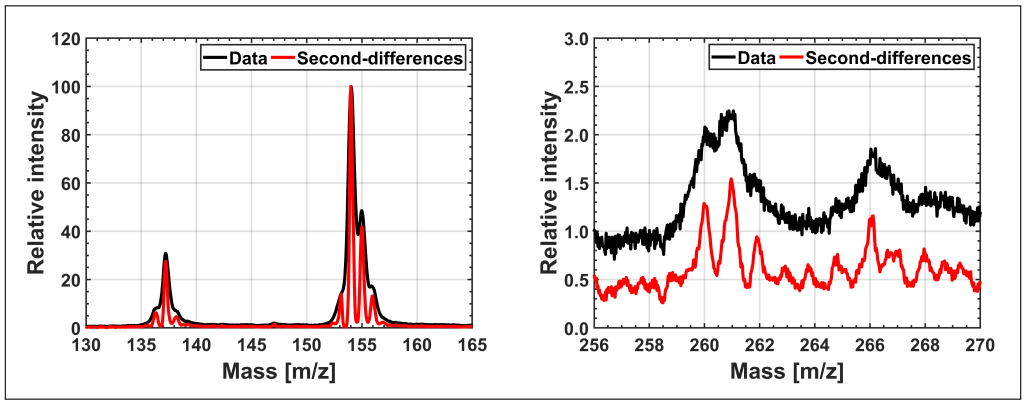



Figure 3: Comparison spectra of the DHB (m/z 154, left) and its dehydroxylated fragment (m/z 137, left), as well the triphenylphosphine (m/z 262, right) ion after deconvolution was performed using the Lucy-Richardson algorithm with the second-differences prior.

## Numerical challenges

We will briefly discuss our algorithm, its usage and arising challenges. It requires the user to make choices regarding the peak shape, the prior distribution, the type of statistics applied, either Poisson (Lucy-Richardson) or Gaussian (ISRA), and a stopping criterion. The choices primarily depend on the spectra being 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 address these choices as they arise. The pseudocode for the algorithm and its variables are discussed in the supporting information.

The primary input into the algorithm consists of the initial values for the signal in the `Signal` vector and the data in the `Data` vector, which are equal in length. Other inputs are two smaller vectors `Prior` and `Peak` which both represent convolution kernels, i.e. representations of the Toeplitz matrices $\mathbb{A}$ and $\mathbb{B}$, which will be convolved with the `Signal` vector. `Signal` is iteratively updated until the convergence criteria are satisfied, while `Data`, `Peak` and `Prior` remain unchanged. The update depends on the convolution of the current `Signal` with the `Peak` and `Prior` vectors. `Peak` was chosen as a symmetric second-order polynomial since the results do not strongly depend on the peak shape, given that it is unimodal.

While the convergence of the entropy $S$ could be used as a stopping criterion, this proved difficult to generalize since its behaviour depends strongly on the given prior. Fortunately, the distribution of the residuals is discussed in the deconvolution literature, which asserts that, for a meaningful reconstruction, the residuals should follow a Gaussian distribution centred around zero.[26] As an example, the distribution of residuals is shown in Figure 4(a) for the second-differences prior. Since this distribution satisfies the above-mentioned requirements, we decided to survey its behaviour to monitor the state of the deconvolution process. Our stopping criterion requires that the *difference in the mean of the residuals* $\Delta\varepsilon$ must decrease below a predefined tolerance $\varepsilon_{\mathrm{TOL}}$ for a predefined number of iterations $\mathrm{N_{MAX}}$. In addition, a constant-sign criterion is implemented due to the oscillation of the *mean of the residuals* $\varepsilon$ around zero (not shown here). Importantly, since the magnitude of the *mean of the residuals* $\varepsilon$ depends on the dataset and prior, it cannot be employed as a universal stopping criterion; rather, the change $\Delta\varepsilon$ should be used. The number of additional iterations to perform once the *difference in the mean of the residuals* has fallen below the threshold (and the oscillations have levelled off) was empirically determined to be $\mathrm{N_{MAX}} = 10$. The respective convergence behaviour of the standard Lucy-Richardson algorithm ($\mathbb{B}_0$) and the second-differences prior ($\mathbb{B}_3$) are compared in Figure 4(b). We confirmed that these trends agree qualitatively with the behaviour of the entropy convergence plots.

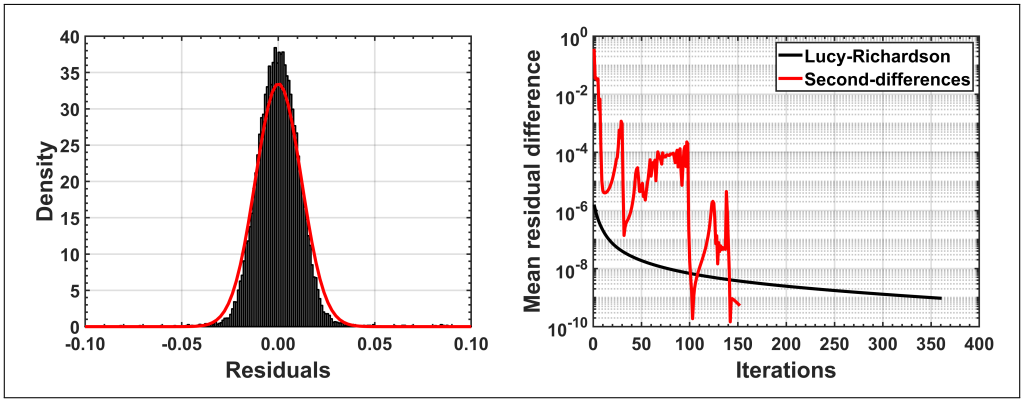



Figure 4: Residual distribution (left) for the second-differences prior discussed in Figure 2. The *difference in the mean of residuals* $\Delta\varepsilon$ are also shown (right) for this prior and standard Lucy-Richardson. The prior reaches the same level of difference $\Delta\varepsilon$ as Lucy-Richardson, but in fewer iterations, which therefore preserves the quality of the reconstruction.

The differences in entropies $\Delta$S will decrease upon approaching local and global minima. Should this difference be below a defined tolerance $\Delta\mathrm{S_{TOL}}$, the strength of the prior is reduced by using scaling the regularization parameter $\beta$ with the amount $\beta_\Delta$ smaller than unity. The attenuation of the prior is akin to the boost mechanism introduced by Miroslav to dislodge the solution out of a local minimum.[66] Figure 5(b) indicates that the prior guides the algorithm during the initial phase of the deconvolution since the influence of $\beta$ is rather significant, but it is increasingly suppressed by the boosting mechanism. Further, note that without such a

boosting mechanism, the Tikhonov regularization prior ($\mathbb{B}_1$) would have been permanently stranded within the local minimum,[67] as indicated by the stepping behaviour of the entropy in Figure 5(a). Most importantly, this mechanism ensures that all the methods converge to the same final entropy, thereby permitting a quantitative comparison of results as a function of different priors for the given dataset, providing an additional reassurance that the selected stopping criterion is indeed appropriate.

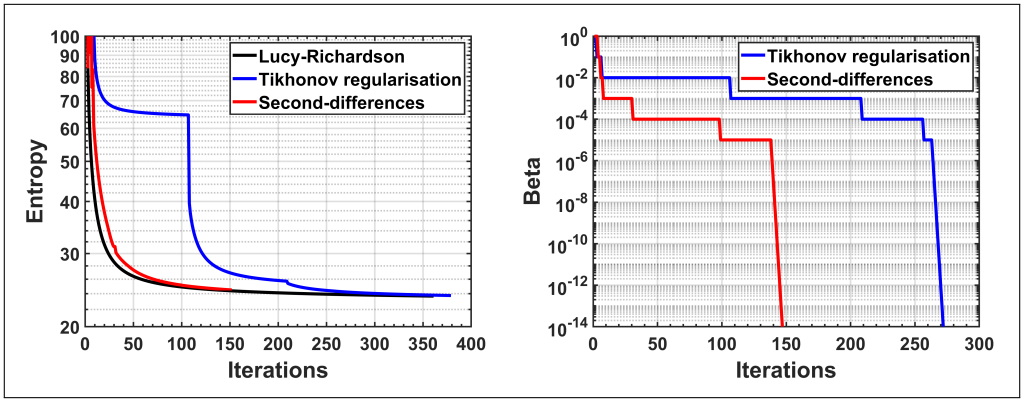



Figure 5: Convergence of the entropy (left) for the standard Lucy-Richardson algorithm and the second-differences and Tikhonov regularization priors. During initial iterations the influence of the prior (right) is large, but it is rapidly suppressed by the boost mechanism.

We emphasize, however, that inclusion of a prior is not beneficial by default. Relative to the standard Lucy-Richardson algorithm, only the second-differences prior improved the rate of convergence (152 versus 361 iterations). The Tikhonov regularization prior produced the highest quality spectra in avoiding overfitting (spurious small peaks) and correctly identifying most of the physically meaningful peaks, but at the expense of the slightly longer convergence time than the Lucy-Richardson algorithm (379 versus 361 iterations). Both the first ($\mathbb{B}_2$) and third-differences ($\mathbb{B}_4$) priors resulted in ill-defined peak shapes, which is not unexpected since these priors discourage sharp features (dependence on signal differences) in the solution.

Although ion counting was performed, it was not known which error distribution would be the most suitable for this data since the mass average is considered in this analysis. For that reason, both Poisson and Gaussian error statistics were tested by using the Lucy-Richardson and the ISRA algorithms respectively. As noted by others, and as also shown in Figure S2, the Lucy-Richardson algorithm converges significantly faster (361 versus 1988 iterations) while providing the superior deconvolution results (the ISRA algorithm is more susceptible to introducing artifact peaks). Currently, it is not known whether this is due to the stopping criterion, although the residual distributions appear to be qualitatively similar, or whether this conclusion applies universally to these algorithms. Even though the results of not using priors are shown, these characteristics were observed for all of the priors investigated. In addition, preliminary results indicate that this observation also applies to single-shot spectra.

# Conclusions

We have demonstrated the advantages of deconvolving low-resolution mass spectrometry data by using the well-established Lucy-Richardson algorithm both with and without priors. Various 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 an accelerated convergence and a preservation of a larger fraction of the peak amplitudes ratios. A novel stopping criterion which monitors the difference in the mean of the residuals was included to facilitate these comparisons, which in combination with a boosting mechanism, ensures that the algorithm does not wind up in a local minimum and that all of the methods reach the same result. The Image Space Reconstruction algorithm was also studied, as it was not known initially whether Gaussian statistics might be more appropriate for the experimental test data. However, as noted previously, Lucy-Richardson converges faster and is less prone to overfitting. For all of the investigated methods, the improved resolution of the deconvolved signal allowed a more precise statement to be made regarding the fragmentation of the benzyltriphenylphosphonium thermometer ion upon femtosecond desorption. Further studies will extend the framework introduced in this work to assist in the interpretation and decoupling of MALDI shot-to-shot phenomena.

# Acknowledgements

Cornelius L. Pieterse and Michiel B. Kock have contributed equally. We 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. R. J. Dwayne Miller is the author of a patent (US8110794B2) related to the mechanism of picosecond infrared laser ablation.

# Supporting Information Available

The algorithm requires the arrays $Signal, Peak, Prior$ and $Data$, representing respectively the initial signal, the peak and prior point spread functions (PSF), and the raw data. With reference to algorithm 1, the $\star$ operator returns the convolution of two arrays, `mean()` returns the mean of the input array, `sum()` the total of an array, `flip()` reverses the elements of an array, and the plus $(+)$ and minus $(-)$ subscripts returns an array containing only positive or negative elements.

**Algorithm 1** Split-gradient method

$N \leftarrow 0; \beta \leftarrow 1.0; \beta_\Delta \leftarrow 0.9; \varepsilon_{tol} \leftarrow 10^{-9}; N_{max} \leftarrow 100; \Delta S_{tol} \leftarrow 0.01$
$(OldEntropy, Recon) = \textsc{GetEntropy}(Data, Signal, Peak)$
$\varepsilon_{old} = \textsc{Mean}(Data - Recon)$
**while** $N < N_{max}$ **do**
    $Conv \leftarrow Signal \star Peak$
    $Penalty \leftarrow Signal \star Prior$
    $\Lambda \leftarrow \textsc{mean}(1 + Penalty^2)$
    $Deriv \leftarrow (Data/Conv) \star \textsc{Flip}(Peak)$
    $Penalty \leftarrow Penalty \star \textsc{Flip}(Prior)$
    $Penalty_+ \leftarrow (Penalty/\Lambda)_+$
    $Penalty_- \leftarrow (Penalty/\Lambda)_-$
    $Signal \leftarrow Signal \times (Deriv - \beta \times Penalty_-)/(1 + \beta \times Penalty_+)$
    $(NewEntropy, Recon) \leftarrow \textsc{GetEntropy}(Data, Signal, Peak)$
    $\Delta S \leftarrow NewEntropy - OldEntropy$
    $OldEntropy \leftarrow NewEntropy$
    $\varepsilon_{new} \leftarrow \textsc{Mean}(Recon - Data)$
    **if** $\Delta S < \Delta S_{TOL}$ **then**
        $\beta = \beta \times \beta_\Delta$
    **end if**
    **if** $(\varepsilon_{old} \times \varepsilon_{new} < 0) \parallel ((\varepsilon_{old} - \varepsilon_{new}) > \varepsilon_{TOL})$ **then**
        $N \leftarrow 0$
    **else**
        $N \leftarrow N + 1$
    **end if**
    $\varepsilon_{old} = \varepsilon_{new}$
**end while**
**function** $\textsc{GetEntropy}(Data, Signal, Peak)$
    $Recon \leftarrow Signal \star Peak$
    $Entropy \leftarrow data \times \log(data/recon)$
    $Out \leftarrow \textsc{Sum}(entropy + signal - data)$
    **return** $(Out, Recon)$
**end function**

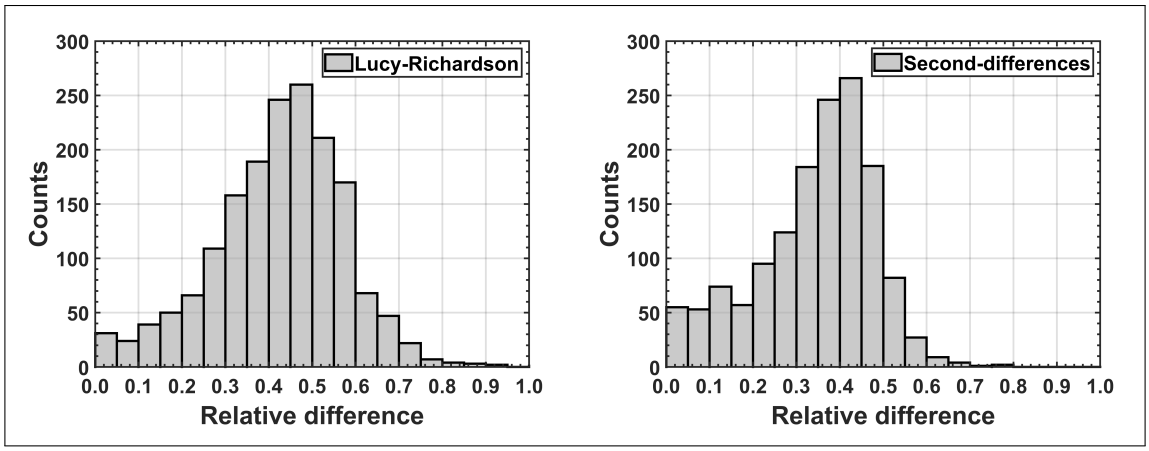



Figure S1: Histograms of the relative differences between the signal and data, with data being the reference, for the Lucy-Richardson (left) and second-differences prior (right) algorithms. Although these distributions have similar centroids, the prior distribution is skewed towards the lower-end, which indicates a better amplitude preservation.

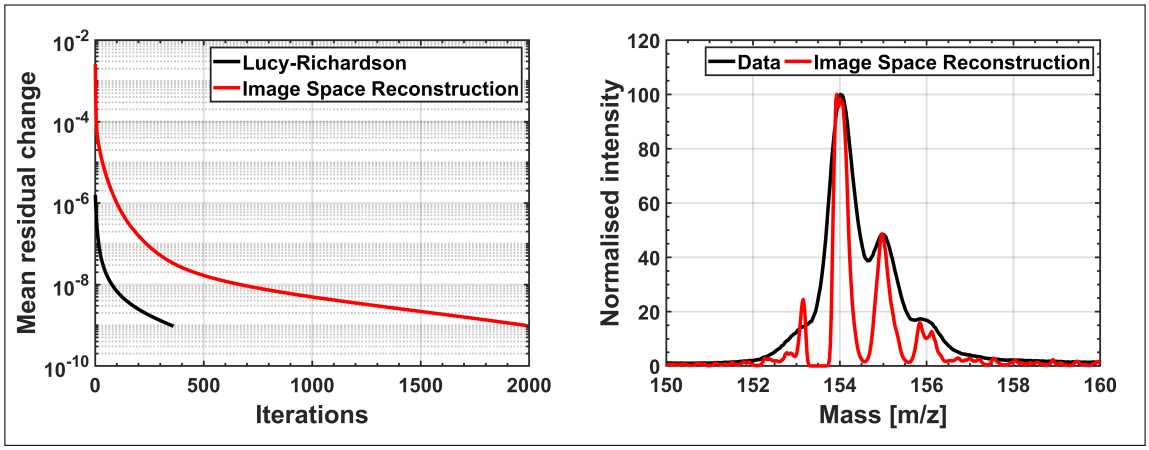



Figure S2: Change in the mean of the residuals (left) for the standard Lucy-Richardson and Image Space Reconstruction algorithms. Also shown are the comparison spectra of the DHB ion (m/z 154, right), clearly showing the introduced artifacts.

# References

(1) Liang, Y.-Z.; Xie, P.; Chan, K. Quality control of herbal medicines. *Journal of Chromatography B* **2004**, *812*, 53–70.

(2) Wiseman, J. M.; Ifa, D. R.; Zhu, Y.; Kissinger, C. B.; Manicke, N. E.; Kissinger, P. T.; Cooks, R. G. Desorption electrospray ionization mass spectrometry: Imaging drugs and metabolites in tissues. *Proceedings of the National Academy of Sciences* **2008**, *105*, 18120–18125.

(3) Ong, S.-E.; Mann, M. Mass spectrometry–based proteomics turns quantitative. *Nature chemical biology* **2005**, *1*, 252–262.

(4) Cravatt, B. F.; Simon, G. M.; Yates Iii, J. R. The biological impact of mass-spectrometry-based proteomics. *Nature* **2007**, *450*, 991–1000.

(5) Diamandis, E. P. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool opportunities and potential limitations. *Molecular & Cellular Proteomics* **2004**, *3*, 367–378.

(6) Diamandis, E. P. Cancer biomarkers: Can we turn recent failures into success? *Journal of the National Cancer Institute* **2010**, *102*, 1462–1467.

(7) Clark, A. E.; Kaleta, E. J.; Arora, A.; Wolk, D. M. Matrix-Assisted laser desorption ionization-time of flight mass spectrometry: A fundamental shift in the routine practice of clinical microbiology. *Clinical Microbiology Reviews* **2013**, *26*, 547–603.

(8) McDonnell, L. A.; Corthals, G. L.; Willems, S. M.; van Remoortere, A.; van Zeijl, R. J.; Deelder, A. M. Peptide and protein imaging mass spectrometry in cancer research. *Journal of proteomics* **2010**, *73*, 1921–1944.

(9) Eberlin, L. S.; Dill, A. L.; Golby, A. J.; Ligon, K. L.; Wiseman, J. M.; Cooks, R. G.; Agar, N. Y. Discrimination of human astrocytoma subtypes by lipid analysis using desorption electrospray ionization imaging mass spectrometry. *Angewandte Chemie International Edition* **2010**, *49*, 5953–5956.

(10) Golf, O.; Strittmatter, N.; Karancsi, T.; Pringle, S. D.; Speller, A. V.; Mroz, A.; Kinross, J. M.; Abbassi-Ghadi, N.; Jones, E. A.; Takats, Z. Rapid evaporative ionization mass spectrometry imaging platform for direct mapping from bulk tissue and bacterial growth media. *Analytical chemistry* **2015**, *87*, 2527–2534.

(11) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. METLIN: a metabolite mass spectral database. *Therapeutic drug monitoring* **2005**, *27*, 747–751.

(12) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E. HMDB 3.0 - the human metabolome database in 2013. *Nucleic acids research* **2012**, *41*, D801–D807.

(13) Fournier, I.; Marinach, C.; Tabet, J.; Bolbach, G. Irradiation effects in MALDI, ablation, ion production, and surface modifications. Part II: 2, 5-dihydroxybenzoic acid monocrystals. *Journal of the American Society for Mass Spectrometry* **2003**, *14*, 893–899.

(14) Wortmann, A.; Pimenova, T.; Alves, S.; Zenobi, R. Investigation of the first shot phenomenon in MALDI mass spectrometry of protein complexes. *Analyst* **2007**, *132*, 199–207.

(15) Chaurand, P.; Schriver, K. E.; Caprioli, R. M. Instrument design and characterization for high resolution MALDI-MS imaging of tissue sections. *Journal of Mass Spectrometry* **2007**, *42*, 476–489.

(16) Bae, Y. J.; Park, K. M.; Kim, M. S. Reproducibility of temperature-selected mass spectra in matrix-assisted laser desorption ionization of peptides. *Analytical Chemistry* **2012**, *84*, 7107–7111.

(17) Ingendoha, A.; Karasa, M.; Hillenkamp, F.; Giessmannb, U. Factors affecting the resolution in matrix-assisted mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes* **1994**, *131*, 345–354.

(18) Dreisewerd, K.; Schürenberg, M.; Karas, M.; Hillenkamp, F. Influence of the laser intensity and spot size on the desorption of molecules and ions in matrix-assisted laser desorption/ionization with a uniform beam profile. *International journal of mass spectrometry and ion processes* **1995**, *141*, 127–148.

(19) Qiao, H.; Spicer, V.; Ens, W. The effect of laser profile, fluence, and spot size on sensitivity in orthogonal-injection matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **2008**, *22*, 2779–2790.

(20) Gilmore, I. S. SIMS of organics - Advances in 2D and 3D imaging and future outlook. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films* **2013**, *31*, 050819.

(21) Anderson, N. L.; Anderson, N. G. The Human Plasma Proteome. *Molecular & Cellular Proteomics* **2002**, *1*, 845–867.

(22) Zhurov, K. O.; Kozhinov, A. N.; Fornelli, L.; Tsybin, Y. O. Distinguishing analyte from noise components in mass spectra of complex samples: Where to cut the noise? *Analytical Chemistry* **2014**, *86*, 3308–3316.

(23) Lu, Y.; Pieterse, C. L.; Robertson, W. D.; Miller, R. J. D. Soft Picosecond Infrared Laser Extraction of Highly Charged Proteins and Peptides from Bulk Liquid Water for Mass Spectrometry. *Analytical Chemistry* **2018**, *90*, 4422–4428.

(24) Skilling, J. *Maximum Entropy and Bayesian Methods: Cambridge, England, 1988*; Springer Science & Business Media, 2013; Vol. 36.

(25) Narayan, R.; Nityananda, R. Maximum entropy image restoration in astronomy. *Annual review of astronomy and astrophysics* **1986**, *24*, 127–170.

(26) Puetter, R.; Gosnell, T.; Yahil, A. Digital image reconstruction: Deblurring and denoising. *Annual Review of Astronomy and Astrophysics* **2005**, *43*.

(27) Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry* **1988**, *60*, 2299–2301.

(28) Dreisewerd, K. The desorption process in MALDI. *Chemical reviews* **2003**, *103*, 395–426.

(29) Bae, Y. J.; Kim, M. S. A thermal mechanism of ion formation in MALDI. *Annual Review of Analytical Chemistry* **2015**, *8*, 41–60.

(30) Knochenmuss, R. The Coupled Chemical and Physical Dynamics Model of MALDI. *Annual Review of Analytical Chemistry* **2016**, *9*, 365–385.

(31) McCrery, D. A.; Peake, D. A.; Gross, M. L. Fast atom bombardment and laser desorption mass spectrometry for determination of alkyltriphenylphosphonium salts. *Analytical Chemistry* **1985**, *57*, 1181–1186.

(32) Claereboudt, J.; Claeys, M.; Geise, H.; Gijbels, R.; Vertes, A. Laser microprobe mass spectrometry of quaternary phosphonium salts: direct versus matrix-assisted laser desorption. *Journal of the American Society for Mass Spectrometry* **1993**, *4*, 798–812.

(33) Claereboudt, J.; Baeten, W.; Geise, H.; Claeys, M. Structural characterization of mono- and bisphosphonium salts by fast atom bombardment mass spectrometry and tandem mass spectrometry. *Journal of Mass Spectrometry* **1993**, *28*, 71–82.

(34) Luo, G.; Marginean, I.; Vertes, A. Internal energy of ions generated by matrix-assisted laser desorption/ionization. *Analytical chemistry* **2002**, *74*, 6185–6190.

(35) Chen, Y.; Vertes, A. Pumping rate and surface morphology dependence of ionization processes in matrix-assisted laser desorption ionization. *The Journal of Physical Chemistry A* **2003**, *107*, 9754–9761.

(36) Luo, G.; Marginean, I.; Ye, L.; Vertes, A. Competing ion decomposition channels in matrix-assisted laser desorption ionization. *The Journal of Physical Chemistry B* **2008**, *112*, 6952–6956.

(37) Mamyrin, B. A.; Karataev, V. I.; Shmikk, D. V.; Zagulin, V. A. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Sov. Phys. - JETP* **1973**, *37*, 45–48.

(38) Brown, R. S.; Lennon, J. J. Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometer. *Analytical chemistry* **1995**, *67*, 1998–2003.

(39) Vestal, M.; Juhasz, P.; Martin, S. Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **1995**, *9*, 1044–1050.

(40) Pieterse, C. L.; Busse, F.; Tellkamp, F.; Robertson, W. D.; Miller, R. J. D. Femtosecond Pumping Rate Dependence of Fragmentation Pathways in Matrix-Assisted Laser Desorption Ionization. *ChemRxiv* **2018**, doi: 10.26434/chemrxiv.6450380.v1.

(41) Strupat, K.; Karas, M.; Hillenkamp, F. 2,5-Dihydroxybenzoic acid: a new matrix for laser desorption–ionization mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes* **1991**, *111*, 89–102.

(42) Lucy, L. B. An iterative technique for the rectification of observed distributions. *The Astronomical Journal* **1974**, *79*, 745.

(43) Richardson, W. H. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America* **1972**, *62*, 55–59.

(44) Bernardo, J. M.; Smith, A. F. *Bayesian theory*; IOP Publishing, 2001.

(45) Shepp, L. A.; Vardi, Y. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging* **1982**, *1*, 113–122.

(46) Geman, S.; Geman, D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1984**, 721–741.

(47) Bertero, M.; Boccacci, P.; Desiderà, G.; Vicidomini, G. Image deblurring with Poisson data: from cells to galaxies. *Inverse Problems* **2009**, *25*, 123006.

(48) Starck, J.-L.; Pantin, E.; Murtagh, F. Deconvolution in astronomy: A review. *Publications of the Astronomical Society of the Pacific* **2002**, *114*, 1051.

(49) Hansen, P. C.; Nagy, J. G.; O'leary, D. P. *Deblurring images: matrices, spectra, and filtering*; Siam, 2006; Vol. 3.

(50) Jansson, P. A. *Deconvolution of images and spectra*; Courier Corporation, 2014.

(51) Young, S.; Driggers, R.; Jacobs, E. Image Deblurring. 2008.

(52) Csiszar, I. Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. *The Annals of Statistics* **1991**, 2032–2066.

(53) Natterer, F.; Wübbeling, F. *Mathematical methods in image reconstruction*; Siam, 2001; Vol. 5.

(54) Barrett, H. H.; Myers, K. J. *Foundations of image science*; John Wiley & Sons, 2013.

(55) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C*; Cambridge University Press, Cambridge, 1996; Vol. 2.

(56) Dempster, A. P.; Laird, N. M.; Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (methodological)* **1977**, 1–38.

(57) Lantéri, H.; Roche, M.; Cuevas, O.; Aime, C. A general method to devise maximum-likelihood signal restoration multiplicative algorithms with non-negativity constraints. *Signal Processing* **2001**, *81*, 945–974.

(58) Lanteri, H.; Roche, M.; Aime, C. Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms. *Inverse problems* **2002**, *18*, 1397.

(59) Daube-Witherspoon, M. E.; Muehllehner, G. An iterative image space reconstruction algorthm suitable for volume ECT. *IEEE Transactions on Medical Imaging* **1986**, *5*, 61–66.

(60) De Pierro, A. R. On the convergence of the iterative image space reconstruction algorithm for volume ECT. *IEEE Transactions on Medical Imaging* **1987**, *6*, 174–175.

(61) Titterington, D. On the iterative image space reconstruction algorthm for ECT. *IEEE Transactions on Medical Imaging* **1987**, *6*, 52–56.

(62) Yang. J. J.,; Gobeli, D. A.; El-Sayed, M. A. Change in the mechanism of laser multiphoton ionization-dissociation in benzaldehyde by changing the laser pulse width. *The Journal of Physical Chemistry* **1985**, *89*, 3426–3429.

(63) Weinkauf, R.; Aicher, P.; Wesley, G.; Grotemeyer, J.; Schlag, E. Femtosecond versus nanosecond multiphoton ionization and dissociation of large molecules. *The Journal of Physical Chemistry* **1994**, *98*, 8381–8391.

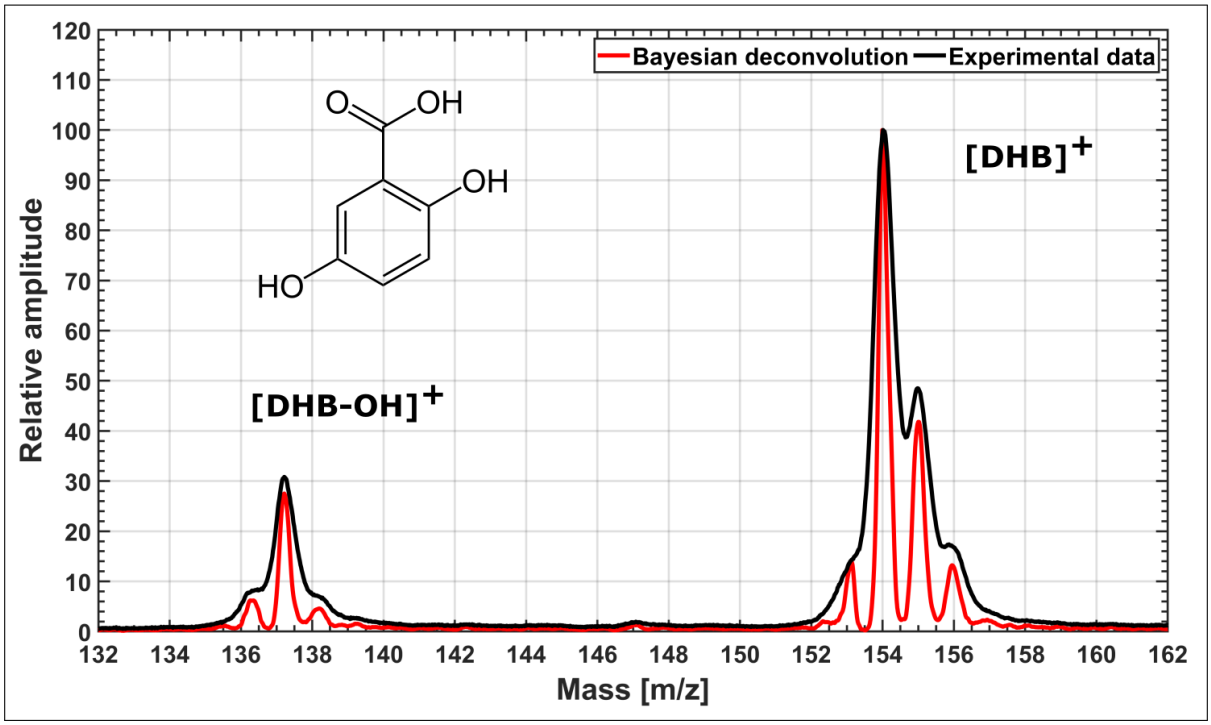
(64) Smith, D.; Ledingham, K.; Kilic, H.; McCanny, T.; Peng, W.; Singhal, R.; Langley, A.; Taday, P.; Kosmidis, C. Ionization and dissociation of benzaldehyde using short intense laser pulses. *The Journal of Physical Chemistry A* **1998**, *102*, 2519–2526.

(65) Markevitch, A. N.; Moore, N. P.; Levis, R. J. The influence of molecular structure on strong field energy coupling and partitioning. *Chemical Physics* **2001**, *267*, 131–140.

(66) Morháč, Miroslav, Deconvolution methods and their applications in the analysis of $\gamma$-ray spectra. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2006**, *559*, 119–123.

(67) Tikhonov, A. N.; Goncharsky, A. V.; Stepanov, V. V.; Yagola, A. G. *Numerical Methods for the Solution of Ill-Posed Problems*; Springer Netherlands: Dordrecht, 1995; p 257.

# For TOC only


Bayesian deconvolution
Experimental data
Relative amplitude
Mass [m/z]
[DHB]+
[DHB-OH]+
O
OH
OH
HO
120
110
100
90
80
70
60
50
40
30
20
10
0
132
134
136
138
140
142
144
146
148
150
152
154
156
158
160
162