Mitigating dead-time effects during multivariate analysis of ToF-SIMS spectral images

Michael R. Keenan,a† Vincent S. Smentkowski,b James A. (Tony) Ohlhausena† and Paul G. Kotulaa†

ToF-SIMS spectra are formed by bombarding a surface with a pulse of primary ions and detecting the resultant ionized surface species using a time-of-flight mass spectrometer. Typically, the detector is a time-to-digital converter. Once an ion is detected using such detectors, the detector becomes insensitive to the arrival of additional ions for a period termed as the (detector) dead-time. Under commonly used ToF-SIMS data acquisition conditions, the time interval over which ions arising from a single chemical species reach the detector is on the order of the detector dead-time. Thus, only the first ion reaching the detector at any given mass is counted. The event registered by the data acquisition system, then, is the arrival of one or more ions at the detector. This behavior causes ToF-SIMS data to violate, in the general case, the assumption of linear additivity that underlies many multivariate statistical analysis techniques. In this article, we show that high-mass-resolution ToF-SIMS spectral-image data follow a generalized linear model, and we propose a data transformation and scaling procedure that enables such data sets to be successfully analyzed using standard methods of multivariate image analysis. Copyright © 2008 John Wiley & Sons, Ltd.

Keywords: dead-time effects; ToF-SIMS; binomial model; weighted PCA

Introduction

Time-of-flight secondary ion mass spectrometry (ToF-SIMS) is a standard tool in the surface analyst’s repertoire. The parallel ion detection afforded by the ToF mass analyzer allows large quantities of spatially registered mass spectra to be acquired in short periods of time. Chemical understanding of such spectral images is facilitated by applying multivariate statistical analysis techniques to extract the essential chemical information from the mountain of raw spectral data.[1–11] Typically, multivariate methods assume that the data can be described by a linear-additive model, that is, the signal intensity arising from a given chemical species at a particular location is assumed to be proportional to the amount of the species present at that site. Owing to detector dead-time effects, however, this assumption is commonly violated when data are acquired under standard conditions (especially when only a few shots/pixel (voxel) are used – as is common in 3D analysis). These effects are frequently exposed through apparent ‘saturation’ of the spectral components. One obvious manifestation is observed when measured isotope ratios differ significantly from natural abundances. Another sign is the near-equal intensities of all the major species in an image. Unfortunately, saturation effects can be easily missed. Therefore it is important to understand and correct for any saturation that might be present automatically.

The origins of detector dead-time effects in ToF-SIMS have been described previously,[12] and can be explained by considering the manner in which ToF-SIMS mass spectra are generated. Secondary ions are produced by bombarding the surface of a sample by a pulse from a primary ion source. These secondary ions are extracted into a ToF mass spectrometer that discriminates ions on the basis of their arrival times at a detector, which, in turn, depend on the ions’ respective mass-to-charge ratios. The mass of a particular ion can then be inferred from the arrival time through a calibration procedure. A complete mass spectrum is built up over time by co-adding the signals from repeated primary ion pulses or ‘shots’, and a spectral image can be formed by rastering the primary ion beam across a sample in one, two or, with interspersed sputtering steps, three spatial dimensions.

For a typical time-to-digital converter used in this application, the detection of a single ion causes the detector to become insensitive to subsequent ion arrivals for a period of time, the duration of which is termed the (detector) dead-time. As is commonly achieved in ToF-SIMS, after a single primary ion pulse most ions derived from the same chemical species arrive at the detector during a time interval that is in the order of the detector dead-time. As a result, only the first ion to arrive registers a count, with subsequent near-simultaneous arrivals being invisible to the detector. This is of little consequence when counting rates are low, i.e. when there is a low probability of detecting an ion at a particular mass on any given shot. In this case, the data follow the linear model and are contaminated with noise characteristic of counting measurements. Adequate ways of treating this type of data in a multivariate sense have been
described previously.\textsuperscript{[13–15]} For high counting rates, however, when the probability of detecting a particular ion on any given shot is high, there is also a high probability that multiple ions will arrive at the detector simultaneously, of which only one will be recorded. This gives rise to nonlinear instrument response and noise characteristics that are non-Poisson. Clearly, as primary ion sources achieve greater efficiency, as with the advent of cluster liquid metal ion sources, such nonlinearity will become a more frequent and prominent feature of ToF-SIMS data sets. New mathematical techniques will be required to effectively extract the chemical information from such spectral data.

The goals of the present work are two-fold. First, using a simple and intuitive example, we will illustrate, in the general case, that ToF-SIMS spectral image data follow a generalized linear model. In other words, the data can be represented as a function of a linear model. Secondly, we propose a simple data preprocessing procedure that can make dead-time-affected ToF-SIMS data amenable to multivariate statistical analysis using standard approaches.

**Statistical Model for Secondary Ion Detection in ToF-SIMS**

The binomial probability distribution plays a fundamental role in understanding each step of the ToF-SIMS measurement process. The statistical properties of secondary ion generation and of ion detection, and the effect of detector dead-time all involve reference to the binomial distribution. In general, the binomial distribution gives the probability $P$ of observing $k$ successes in $n$ trials, given the probability $s$ for success on any single attempt. Here, success may mean the desorption of a particle from many like particles bombarded by a primary ion pulse, the actual detection of a species-specific secondary ion from several that might be produced by the ionization event, or the observation of one or more such secondary ions following one primary ion pulse in a series of pulses. The mathematical formulation of the binomial probability distribution is\textsuperscript{[16]}:

$$f(k; n, s) = \frac{n!}{k!(n-k)!} s^k (1-s)^{n-k}, \quad k = 0, 1, 2, \ldots, n$$

The mean $\mu_B$ and variance $\sigma_B^2$ of a binomially distributed random variable are given by:

$$\mu_B = ns$$

$$\sigma_B^2 = ns(1-s)$$

Previously, researchers\textsuperscript{[12]} have noted that if a portion of a sampled surface has $n$ particles of a given type, and probability $s$ that any one particle will desorb during a small interval of time, the probability that $k$ such particles will desorb during the time interval is given by Eqn (1). According to Eqn (2), the mean number of particles desorbing is expected to be $ns$. In other words, the number of desorbing particles follows a linear model; the number desorbed is linearly related to the number sampled. Most commonly, only a few of the surface particles will be desorbed on any single pulse, that is, the probability of desorption $s$ is small. In this limiting case, the binomial distribution can be approximated by the well-known Poisson distribution with parameter $\lambda$\textsuperscript{[16]}:

$$f(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

The mean $\mu_P$ and variance $\sigma_P^2$ of the Poisson distribution are both equal to $\lambda = ns$. Again, the expected number of desorbed particles of any given type follows a linear model.

After the surface species are desorbed, a number of factors influence the probability that they will be ionized and detected by the instrument.\textsuperscript{[17,18]} Lumping these into a single probability $q$, and assuming a perfect, linear detection system (i.e. no noise, and the detection probability is independent of the numbers of ions both generated and arriving at the detector), the probability of counting $d$ ions given that $k$ particles of that type were desorbed is given by the binomial distribution:

$$P(d|k) = \frac{k!}{d!(k-d)!} q^d (1-q)^{k-d}$$

The overall probability of observing $d$ ions is then obtained by summing the conditional probabilities of Eqn (4) over all possible values of $k$, each weighted by the probability that $k$ particles were desorbed:

$$P(d) = \sum_{k=d}^{\infty} \frac{(ns)^k}{k!} e^{-ns} \frac{k!}{d!(k-d)!} q^d (1-q)^{k-d}$$

This sum is readily evaluated to yield the intuitively appealing result:

$$P(d) = \frac{(nsq)^d}{d!} e^{-nsq}$$

That is, the number of ions of a particular species detected will be described by the Poisson distribution with mean $nsq$. Since the expected mean of the observed counts is proportional to $n$, the total number of corresponding particles sampled, the data are seen to be linearly related to the composition of the sample. Note that it is not required to assume that the detection probability is small to arrive at the Poisson distribution.

The foregoing development assumed that the detection system has a constant probability of detecting an ion of a particular type, independent of the rate at which the ions reach the detector. This will generally not be the case. For detectors typically used in ToF-SIMS applications, the detection of one ion causes the detector to become insensitive to the arrival of additional ions for a period of time termed the detector dead-time. Several ways in which the dead time can affect ToF-SIMS spectra have been described previously.\textsuperscript{[12]} In this article, we only consider the common limiting case where all ions of the same species arrive at the detector within the dead time. Additionally, we assume that the masses of all distinct ion species differ sufficiently so that they can be counted independently of one another. Subsequent to a single primary ion pulse, in this scenario, either no ions will be detected at a specific mass, or one event will be observed denoting the arrival of one or more ions at the detector. For a series of $m$ primary ion pulses, the total number of counts $c$ obtained at a particular nominal mass is, again, binomially distributed:

$$f(c; m, p) = \frac{m!}{c!(m-c)!} p^c (1-p)^{m-c}$$

Here, $p$ is the probability an ion event is seen on any single shot. Since it is assumed that the detector is fully sensitive prior to the
Mitigating dead-time effects

detection of the first ion in a group, the required probability is simply the sum over all nonzero \( d \), of the probabilities in Eqn (6):

\[
p = \sum_{d=1}^{\infty} P(d) = 1 - P(0) = 1 - e^{-nsq}
\]

where the latter equalities follow from the fact that the sum of probabilities over all possible events is equal to unity. The expected number of total counts \( C \) for a specific ion given \( m \) total primary ion pulses is, according to Eqn (2):

\[
C = mp = m(1 - e^{-nsq})
\]

Additionally, the variance is given by

\[
\sigma_C^2 = mp(1 - p) = Ce^{-nsq}
\]

It is notable that, in the general case, the expected number of counts is no longer directly proportional to the number of analogous particles in the sample; rather, it is an exponential function of \( n \). Of course, if the probability is small that even one ion will be detected on any given shot, the exponential can be approximated by a linear term and the number of counts becomes Poisson-distributed with parameter \( C = nsq \), and the linear model is recovered. Equation (9) can serve as the basis for a dead-time correction procedure.\(^{12}\) Rearranging this equation, taking the logarithm and multiplying both sides by \( m \) yields:

\[
nsq = -m \ln \left( 1 - \frac{C}{m} \right) \equiv \tilde{C}
\]

Thus, \( n \) becomes linearly proportional to a transformed representation of the data \( \tilde{C} \). The variance of the transformed data, which will be useful later, can be approximated as\(^{16}\):

\[
\sigma_{\tilde{C}}^2 = \frac{C}{1 - \frac{C}{m}}
\]

The effectiveness of this approach for a specific mass depends on having a significant number of primary ion pulses for which there are no detected ions at that mass. Note that as it becomes more likely that such a secondary ion is detected on every pulse, \( \tilde{C} \) approaches infinity.

**Multivariate Image Analysis**

Standard approaches to dead-time correction require aggregating secondary ions arising from a large number of primary ion pulses to compose each spectrum. For spectral images, which may contain thousands to millions of individual spectra, taking many looks at each location is frequently not practical. A recently described 3-spatial-dimension spectral-image data set\(^{8}\) for instance, contained approximately 4 million spectra. Acquisition-time constraints allowed only 5 shots per voxel to build up each spectrum. Fortunately, the high level of redundancy that is typically present in spectral images can be exploited to mitigate dead-time effects through use of multivariate statistical techniques.

Multivariate methods such as principal component analysis (PCA) and multivariate curve resolution (MCR) have been applied to ToF-SIMS spectral images.\(^{2,4,8,14}\) A variety of other multivariate approaches that are suitable for ToF-SIMS in the low-count-rate (Poisson) regime have also been described recently.\(^{13}\) The basic strategy shared by all of these methods is to factor a large matrix containing the mass spectral data into a set of smaller component matrices that capture the chemical information present in the data set. If \( \mathbf{D} \) is a \( u \times v \) matrix consisting of a \( v \)-channel mass spectrum at each of \( u \) spatial locations, a typical factorization is:

\[
\mathbf{D} = \mathbf{AS}^T + \mathbf{E}
\]

Here, \( \mathbf{S} \) is a \( v \times c \) matrix that describes the spectral characteristics of the \( c \) components comprising the sample, and \( \mathbf{A} \) is an \( u \times c \) matrix showing their spatial distributions, or abundances. A component, in this case, can be a single chemical species or a combination of species that vary together in the same way. \( \mathbf{E} \) is a matrix of random errors assumed to have zero mean. Equation (13) is an instance of a linear model; the data are represented as a weighted sum of component abundances. In the presence of detector dead-time effects, however, an analog of Eqn (9) provides a more realistic model of the data:

\[
\mathbf{D} = m(1 - e^{-\mathbf{AS}^T}) + \mathbf{E}
\]

where \( \mathbf{1} \) is a \( u \times v \) matrix of ones and the exponential is computed element-wise. The matrix defined by the term in parentheses represents a set of probabilities, one for each element of \( \mathbf{D} \). Assuming, again, zero-mean errors, the abundance and spectral matrices \( \mathbf{A} \) and \( \mathbf{S} \) can be estimated by first transforming \( \mathbf{D} \) as in Eqn (11) above:

\[
-m \ln \left( 1 - \frac{\mathbf{D}}{m} \right) \equiv \tilde{\mathbf{D}} = \mathbf{AS}^T
\]

and then applying standard methods of factor analysis to the transformed data \( \tilde{\mathbf{D}} \). Similarly, the estimated element-wise variance \( \tilde{\mathbf{V}} \) of the transformed data is given by an expression analogous to Eqn (12)

\[
\tilde{\mathbf{V}} = \frac{\mathbf{D}}{1 - \frac{\mathbf{D}}{m}}
\]

Two problems remain. First, for highly saturated data, elements of \( \mathbf{D} \) may equal the number of shots per spectrum \( m \), producing infinities in the transformed data and estimated variance. Second, the noise in the transformed data is not uniform, that is, the magnitude of the uncertainty in a given element of \( \tilde{\mathbf{D}} \) depends on the corresponding number of counts. The first problem can be addressed in a number of ways. The approach employed here is to replace elements \( \bar{d} \) of \( \mathbf{D} \) that equal \( m \) with a mean value \( \bar{d} \) estimated from similar pixels. Specifically, for data in a given mass channel that is binomially distributed, the ratio \( R \) of the number of elements with \( d = m \) to the number having \( d = m - 1 \) is

\[
R = \frac{\bar{d}}{m} \frac{m}{1 - \frac{\bar{d}}{m}}
\]

The desired value of \( \bar{d} \) is then found to be

\[
\bar{d} = \frac{m^2R}{1 + mR}
\]

Note that if \( R < 1, \bar{d} < m - 1 \), and we replace \( d \) with \( m - 1 \) in these cases. The process is repeated for all mass channels containing fully saturated pixels.
For Poisson-distributed data, a data-scaling method has proven very effective at accounting for nonuniform noise while performing multivariate statistical analyses.\textsuperscript{14,15} In this approach, the data matrix $D$ is pre- and postmultiplied by diagonal scaling matrices $G$ and $H$, respectively:

$$D_s = GDH$$ \hspace{1cm} (19)

The scaled data matrix $D_s$ is then factored using standard methods

$$D_s = AS_iS_i^T$$ \hspace{1cm} (20)

and the component matrices are recovered by applying the inverse scaling to the factors:

$$A = G^{-1}A_S$$

$$S = H^{-1}S_S$$ \hspace{1cm} (21)

A derivation\textsuperscript{14} of the optimal scale factors in the case of Poisson data shows that the diagonal of $G$ is simply computed as the inverse square root of the row means of $D$ and that the diagonal of $H$ is computed as the inverse square root of the column means. The derivation makes use of the fact that for Poisson data, the matrix of estimated variances $V$ is simply equal to $D$ itself. A similar approach can be used to account for nonuniform noise in the binomial case. The only difference is that the estimated variance matrix $V$ for the transformed data given by Eqn (16) must be used in place of $D$ to compute the scaling matrices. To summarize, given a $u$-vector $V_{row}$ consisting of the $u$ row means of $V$ and a $v$-vector $V_{col}$ containing the $v$ column means:

$$G = \text{diag} \left( \frac{1}{V_{row}} \right)$$

$$H = \text{diag} \left( \frac{1}{V_{col}} \right)$$

$$\tilde{D}_s = GDH$$ \hspace{1cm} (22)

The transformed and scaled data matrix can then be factored in a manner analogous to Eqn (20), and to the physically meaningful components obtained by inverse transformation as in Eqn (21).

\section*{Experimental}

The sample employed in this work is a diffusion couple that was fabricated from high-purity copper and nickel disks. Details of the sample’s construction, as well as, a multivariate analysis of the sample by energy-dispersive X-ray spectroscopy have been presented earlier.\textsuperscript{19} An optical micrograph of the sample and of the interface that was imaged is shown in Fig. 1. The voiding evident at the interface is a consequence of the diffusion process. The sample was prepared using standard metallographic techniques; thus, the voids presented opportunities to trap residues from the polishing compound. ToF-SIMS spectral images were acquired using the high-current bunched pulsing mode of the Bi-ion source on the ION-ToF model ToF-SIMS IV instrument. The region analyzed was sputter cleaned \textit{in situ} using 3 kV Ar ions. Each spectral image was 350 micron $\times$ 350 micron in size with a pixel density of 128 $\times$ 128, and was obtained from the same area as represented by the box in Fig. 1. Every image was composed from 100 raster frames with one primary ion shot used per pixel per frame. Mass spectra were compiled by binning the ions detected in the range from 1 to 100 amu in 1/2 amu increments (199 total mass bins). The mass-bin edges were placed at half-integral and integral masses, thus, each integral mass comprised two mass bins. For an integral mass $M$, the two bins were denoted $M-$ and $M+$ to represent the bins containing masses less than and greater than the nominal mass, respectively. The intent here, was to separate ions arising from metal species from hydrocarbon ions at the same nominal mass. Finally, to conform to the assumption inherent in the dead-time correction model, only one ion per mass per shot was allowed into the data set, even if more than that were present in the raw data file. As will be detailed later, this was a minor adjustment.

A total of 14 spectral images were analyzed. The high-current bunched pulsing mode provides for the highest sensitivity and is prone to peak saturation resulting from detector dead-time effects. In the high-current bunched mode settings file, parameters such as the primary ion peak width and suppressor were adjusted to vary the emission current and hence the number of secondary ions produced. Five of the spectral images, which span the range of observed ion intensities, were selected for multivariate analysis. For statistical purposes, three regions of interest (ROI) were defined and intended to represent pure copper, pure nickel and the intermixed diffusion zone. The ROIs are outlined by the boxes superimposed on the total ion image in Fig. 2. The ROIs were selected as being regions showing relatively constant signal intensity. By minimizing the contributions due to spatial variation, the statistical fluctuations associated with the measurement process itself are more easily evaluated. Fortunately, being images, large numbers of samples are available allowing good statistical estimates to be made. The ROIs comprise 1800, 1425 and 1440 pixels (mass spectra), respectively, for Cu, Ni and the mixed region. Several characteristics of the mass spectral data in the copper ROI are summarized in Table 1 for the five data sets. Zinc, which is present as a minor impurity in the copper, was used, as will be outlined later, to provide an indirect indication of the relative extents of secondary ion generation. The measured $^{64}$Zn ion intensity in the copper ROI varied by a factor of 27 over the range of acquisition conditions employed here, suggesting that secondary ion production spanned at least that range.
Mitigating dead-time effects

Figure 2. Regions of interest (ROI) defined to represent (a) pure Cu, (c) pure Ni and (b) the diffusion zone in the total ion image. The mean mass spectra over each of these ROIs in the vicinity of the Cu- and Ni-isotope masses are also shown for the data set with the highest count rate (Dataset E in Table 1). The isotopic ratios are drastically different than accepted natural abundance values.

Table 1. Several metrics characterizing the signal variation observed in the copper ROI for five spectral images that span the full range of observed ion intensities

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Mean (^{63}\text{Zn counts})</th>
<th>Mean (^{63}\text{Cu counts})</th>
<th>% Saturated (^{63}\text{Cu pixels})</th>
<th>Measured (^{63}\text{Cu} : ^{65}\text{Cu})</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.12</td>
<td>20.0</td>
<td>0.0</td>
<td>2.15</td>
</tr>
<tr>
<td>B</td>
<td>0.42</td>
<td>51.0</td>
<td>0.0</td>
<td>1.88</td>
</tr>
<tr>
<td>C</td>
<td>0.98</td>
<td>80.3</td>
<td>0.0</td>
<td>1.58</td>
</tr>
<tr>
<td>D</td>
<td>2.10</td>
<td>97.4</td>
<td>7.1</td>
<td>1.22</td>
</tr>
<tr>
<td>E</td>
<td>3.34</td>
<td>99.8</td>
<td>85.4</td>
<td>1.06</td>
</tr>
</tbody>
</table>

In order to establish ‘ground truth’, traditional retrospective analyses were performed on each of the five selected images in Table 1 by stepping through the entire set of 199 individual ion maps. Depending on the acquisition conditions, between three and six components could be readily identified. Figure 3 shows six representative ion maps from Dataset D, the data set collected with the greatest apparent sensitivity. Copper and nickel are, as expected, the most prevalent species. In addition, adventitious hydrocarbons are seen throughout the image with some perceptible enhancement at the interface, and alkali contamination (notably Na) is observed, primarily associated with copper region. Two isolated components are also found in the diffusion zone; one has mass peaks characteristic of chromium, and the other has a series of peaks at 31, 45, 59 and 73 amu, possibly \(\text{C}_n\text{H}_{2n+1}\)O species. While these two components are situated close to one another, they are not spatially co-located.

All calculations presented here were performed in the Matlab environment (Release 2006b, The Mathworks, Inc., Natick, MA, USA) using custom routines developed in-house.

Results and Discussion

The spectral image shown in Fig. 2 represents the data set acquired at the highest counting rate (Dataset E). The mean spectra obtained from each of the three ROIs are also shown and the effects of detector dead-time are plainly visible. Most telling are the apparent \(^{63}\text{Cu} : ^{65}\text{Cu}\) and \(^{58}\text{Ni} : ^{60}\text{Ni}\) isotope ratios. In each case, the observed isotope ratio is approaching unity. This contrasts with the accepted natural-abundance ratios\(^{[20]}\) of 2.24 and 2.60 for copper and nickel, respectively. In the diffusion zone, the isotopic ratios are somewhat larger as might be expected since the concentrations of both elements are lessened in this region, which should give rise to lower counting rates for each. Interestingly, the total ion yield in the mixed region appears to be enhanced by about 30% over that in the pure-element areas. This is largely an artifact resulting from the detector dead-time. Dead-time effects cause ion detection at a particular mass to be increasingly impeded as the rate of ion generation grows. These effects do not inhibit detection of ions at sufficiently different nominal masses, however. Thus, detection of Ni and Cu are separately less inhibited leading to a sum of the two species that is larger.

Additionally, it is notable that mean number of counts is close to 100 for \(^{63}\text{Cu}\) and \(^{58}\text{Ni}\); 100 primary ion pulses were used to generate each spectrum. In the limit that a detector becomes completely insensitive to all ions of single species after the first is counted, no more than \(m\) counts should be observed at any nominal mass where \(m\) is the number of ‘shots’ per spectrum. Figure 4 shows histograms for the ions detected in the \(^{63}\text{Cu}\) and
65Cu mass channels as acquired. In this extreme case, only 2 pixels out of 16,384 had as many as 102 counts in the 100 total frames, and fewer than 2% of the pixels had 101 counts in the raw data. It would appear that the dead-time limit is largely achieved, and it was enforced here by strictly limiting the number of counts per channel to one per shot for the multivariate analysis. Finally, it is commonly assumed that counting data follows the Poisson distribution. In that case, the expected variance of the data is equal to the mean. Equivalently, the standard deviation data should be approximately equal to the square root of the mean. For about 100 counts, one would expect a standard deviation of about 10. Clearly, the widths of the histogram peaks for both 63Cu and 65Cu in Fig. 4 exhibit much reduced variation as compared to the Poisson model. At low counting rates, on the other hand, the data follow the Poisson model. Zinc is present as a minor impurity in the copper. The histogram of 64Zn counts detected in the Cu ROI is shown in Fig. 5 for Dataset E. The mean number of 64Zn ions is 3.34 with a variance of 3.24, and the distribution of counts closely follows the Poisson prediction.

PCA was chosen here to be the exemplary multivariate statistical method; however, the same principles apply to other techniques such as MCR-ALS. Three different data pretreatment methods were examined. In the first, it was assumed that the data follow the Poisson model, and the standard Poisson scaling was applied. The second approach corrected the data for dead-time effects. The observed value of 2.25 is in excellent agreement with the accepted value of 2.24. Given the probability of detecting a given ion, i.e., $C/m$ in Eqsns (9) through (12), Eqn (11) can be used to estimate the number of counts that would have been observed in the absence of dead-time effects. This prediction is shown for 63Cu on the top scale of Fig. 7. For the most saturated data, with the binomial probability of detecting an ion approaching one, five detectable ions are lost for every one detected.

The second approach corrected the data for dead-time effects...
Mitigating dead-time effects

Figure 7. Binomial probabilities of observing copper-isotope ions as a function of a surrogate measure of ion yield. The curves are fits of the respective data to the shot-normalized version of Eqn (9).

According to Eqn (11) and the ensuing discussion, this is similar to the so-called Poisson correction available in some commercial ToF-SIMS analysis software. After applying the correction, the data were treated as if they were Poisson-distributed. Thirdly, binomial scaling was applied to the corrected data as outlined in Eqn (22). Principal components were computed using the kernel PCA approach. After applying the inverse scaling to the principal components, the components were re-orthogonalized using the factor-PCA algorithm presented previously.

The purpose of data scaling is to make the noise appear more uniform, that is, a constant independent of the signal’s magnitude. For uniform noise, the sorted eigenvalues of the data cross-product matrix \( D^TD \) can be used to infer how many sources of spectral variation are giving rise to the informative part of the data. Briefly, on a semilog plot, eigenvalues corresponding to components describing noise fall on a relatively horizontal line. As systematic, chemically induced signal begins to add to the noise, the eigenvalues rise above this line; thus, the number of significant components can be estimated by looking for the break in the curve. Figure 8 compares the eigenvalue plots subsequent to the three different data pretreatment procedures. Well-defined noise floors are present in all cases, but the apparent number of significant components varies widely with both overall ion intensity and data pretreatment. Table 2 lists the number of significant components estimated from these plots, although it must be noted that this number is not always unambiguous.

Score images and loading vectors corresponding to the first principal component are shown in Fig. 9 for the five data sets and three different analysis strategies. This component mainly describes the copper distribution; a hydrocarbon signature and a small amount of sodium are also present. For Dataset A, all three methods give essentially the same result. This might be expected since that spectral image has the lowest counting rate and is least perturbed by dead-time effects. The correction procedure causes little change to the data and the Poisson approximation holds reasonably well. As the degree of secondary ion generation increases, the three analysis approaches yield distinctly different results. Since the Poisson model does not account for lost signal at high counting rates we see two effects. First, as the \( ^{63}\text{Cu} \) ion intensity will be relatively more inhibited than the \( ^{65}\text{Cu} \) intensity, we observe significant changes in the copper-isotope ratio. Second, since both copper isotopes will show reduced intensity with respect to minor spectral features, the level of hydrocarbon contamination appears to rise dramatically with increased intensity. We also note that the crispness of the interface fades in the score images as copper becomes relatively less important for describing the overall signal in this component.

The two remaining methods correct the observed ion count to account for detector dead-time. In these cases, the loading vectors look similar, and are essentially constant across all levels of ion intensity. When the data are scaled according the Poisson model, however, the score images take on a more speckled appearance with increasing intensity. The reason for this is that while the

Figure 8. Eigenvalues of the data cross-product matrices for the five data sets of Table 1 under three different assumptions concerning the statistical distribution of the data. In (a) the original data are assumed to follow the Poisson model. The data are corrected in (b) according to Eqn (11) but are still treated as Poisson. The binomial model, including both correction and proper scaling, is used for (c).
Table 2. Summary of multivariate analysis results for five spectral images that span the full range of observed ion intensities. The gray areas indicate principal components that largely describe noise in the data, not real chemical features.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Significant eigenvalues</th>
<th>Principal component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
</tr>
<tr>
<td>Poisson model (original data)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>3 Cu Ni Cr</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>4 Cu Ni Na Cr</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>5 Cu Ni Na Cr</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>6 Cu Ni Na Cr CHO Cr</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>5 Cu Ni Na Cr</td>
<td></td>
</tr>
<tr>
<td>Poisson model (corrected data)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>3 Cu Ni Cr</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>5 Cu Ni Na Cr CHO Cr</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>7 Cu Ni Na Cr</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>9 Cu Ni Na Cr CHO Cr</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>8 Cu Ni Na Cr</td>
<td></td>
</tr>
<tr>
<td>Binomial model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>3 Cu Ni Cr</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>4 Cu Ni Na Cr</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>4 Cu Ni Na Cr CHO Cr</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>5 Cu Ni Na Cr CHO Cr</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>4 Cu Ni Na Cr</td>
<td></td>
</tr>
</tbody>
</table>

Figure 9. Score images and loading vectors corresponding to the first principal component for the five data sets of Table 2, and three different data pretreatment methods. This component largely describes the copper distribution in the spectral image.

correction procedure is designed to yield the correct mean, it does not add any noise, so the variances will be incorrect. To take one example, 98 counts will always be corrected to 391, and 99 counts will always be corrected to 461 if there are 100 shots per spectrum (see Eqn (11)). Since there is no variation about these points, the intensities become quantized in large steps giving rise to the speckled appearance. Properly scaling the corrected data to account for the variance predicted by the binomial model gives rise to both loading vectors and score images that are largely independent of the data acquisition conditions. Thus the interpretation of the data should be relatively less subject to purely instrumental effects. Note that there is some loss of
Mitigating dead-time effects

spatial resolution at higher ion intensities. This loss of resolution is apparent in the raw data and was not introduced by the statistical analysis.

The second principal component mainly describes the nickel phase, and the score images and loading vectors exhibit trends that are qualitatively similar to the copper component. They do not provide any additional insights. Things become more interesting when the minor chemical species in the subsequent principal components are considered. Figure 10 shows the third principal component for the same 15 combinations of images and processing. Again, Dataset A meets the assumptions of the Poisson model and gives the same result for all three analysis approaches. The component found here is the chromium inclusion. For the remaining data sets analyzed by correcting the data but applying the Poisson-model scaling, the third most important component primarily describes noise in the copper and nickel isotopes. That is, this noise is found more important than a real chemical component for describing the overall variation in the data set. The third principal component for Dataset B is found to describe sodium and enriched interfacial hydrocarbons using both the Poisson and binomial models. For the hydrocarbons, the enrichment is inferred from the negative quantity of a negative component, and is consistent with the results of the retrospective analysis. As the counting rates increase, however, the Poisson model starts describing variation in the Ni and Cu isotope ratios. The effect is most pronounced at the interface where, as discussed previously, the major isotope is less suppressed relative to the minor isotope for each of these elements. Interestingly, the hydrocarbon signature is still present, but the positive score at the interface combined with a negative hydrocarbon profile suggests that hydrocarbons are depleted at the interface, contrary to the known composition. Using the correction and scaling appropriate to the binomial model, the score images and loading vectors for Datasets B–E are, again, quite similar and are consistent with the known sodium and hydrocarbon distributions.

The remaining principal components tell a similar story. For Dataset A, the one acquired with the lowest ion yield, only three principal components are significant and they have been described above. As the ion yield increases, additional significant components are found. For Dataset D, all six species identified in the retrospective analysis of Fig. 3 are discovered by PCA. Additionally, all data preprocessing methods used here find all of the species if enough principal components are retained in the model. The difference is that principal components describing noise become interspersed among those representing real chemical information when an incorrect statistical model is assumed. While that may not be a major problem here given such a simple sample, severe difficulties for interpretation can arise when the difference between what constitutes signal and noise becomes more obscure. Table 2 summarizes all of the PCA results and indicates which principal components correspond to chemical features, and which describe noise. The results obtained with the binomial model are unique in having no noise components that appear more important than the least significant chemically relevant component. The worst results were obtained for data that was corrected for dead-time effects but not appropriately scaled for binomially distributed noise. This suggests that great care should be taken when applying corrections to ToF-SIMS data within commercial software packages if the data is destined for multivariate analysis.

Figure 10. Score images and loading vectors corresponding to the third principal component for the five data sets of Table 2, and three different data pretreatment methods. HC indicates the hydrocarbon fingerprint.
Summary

We have shown that ToF-SIMS spectral images can be impacted by detector dead-time effects under commonly used data acquisition conditions. In fact, in order to exploit the large dynamic range available to ToF-SIMS, and to be sensitive to minor chemical species, it may be desirable to operate with ion yields that are assured of being dead-time limited. This is particularly true for 3D images where only one or a few primary ion pulses can be applied per voxel. Dead-time effects are manifested in the statistical distributions of the data, which become binomial rather than Poisson at modest to large counting rates. Under these conditions, the spectral data become nonlinear functions of sample composition. This has significant implications for quantification and spectral interpretation. Many analysts use isotopic ratios and/or ‘expected fingerprint patterns’ to positively identify species. This will be especially important for biological analysis where groups of high mass peaks may have a priori expected peak spacings and relative intensities. Recalling the total-ion-enhancement artifact in Fig. 2, we note the more subtle problems that dead-time effects can pose for spectral image interpretation.

Another consequence of the binomial model is that the basic assumption that underlies many multivariate image analysis techniques, namely, that the data follow a linear additive model, can be severely violated. We have also proposed data transformation and scaling method that allows standard multivariate techniques to be employed with binomially distributed data. Excellent, intuitively appealing results were achieved with a simple sample. Finally, this discussion points out the need for analysts to fully understand the statistical nature of their data if they expect to achieve optimal results from multivariate analysis.

Acknowledgements

The authors would like to thank ION-TOF for providing the information required to translate the raw spectral data files. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy’s National Nuclear Security Administration under Contract DE-AC04-94AL85000.

References
