Unfolding the background of secondary ions in measured nanodosimetric ionisation cluster size distributions

G. Hilgers, M.U. Bug and H. Rabus

Physikalisch-Technische Bundesanstalt,
Bundesallee 100, 38116 Braunschweig, Germany

E-mail: gerhard.hilgers@ptb.de

ABSTRACT: Nanodosimetry is a methodology for quantifying the effects of ionising radiation on matter by determining the frequency distributions of the cluster size of ionisations in nanometric target volumes. In previous investigations with the Ion Counter nanodosimeter operated at PTB, significant deviations for large cluster sizes were found in the comparison between measured and simulated data of ionisation cluster size distributions. These deviations could be explained by a background of secondary ions, which are produced within the transport system of the ionised target molecules. In this paper, two different approaches were investigated to correct for the background of secondary ions in the measured data to obtain the "true" cluster size distribution to be used, e.g., in predictions of biological effectiveness. In the first approach, the correction of the background was treated as a minimising problem. In the second approach, an iterative unfolding algorithm using Bayes statistics was employed. In all cases where the convolution of the background-corrected results with the secondary ion background agrees well with the corresponding measured cluster size distribution, the background-correction led to an improved agreement between measurement and simulation. For the removal of a background of secondary ions from measured cluster size distributions, the unfolding algorithm using Bayes statistics is the preferred method as it proved to be the most effective and the least sensitive to boundary conditions. Moreover, it was considerably less time consuming.

KEYWORDS: Data processing methods; Microdosimetry and nanodosimetry

© 2019 The Author(s). Published by IOP Publishing Ltd on behalf of Sissa Medialab. Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
1 Introduction

Nanodosimetry investigates the ionisation component of the particle track structure, which is characterised by the relative frequency distribution of the ionisation cluster size [1]. The ionisation cluster size denotes the number of ionisations \( \nu \) created in a target volume by a primary particle and its secondaries. A primary particle can either traverse the target volume or pass it at a distance \( d \) with respect to its centre. The ionisation cluster size distribution is the statistical distribution of the probabilities \( P_\nu(Q, d, \rho) \) that exactly \( \nu \) ions are created in the target volume for a radiation quality \( Q \), with the primary particles passing the target volume at a distance \( d \) in a target gas of density \( \rho \). Assuming a target volume of cylindrical symmetry and \( d \) being the distance from the central axis of the cylinder, it should be noted that the ionisation cluster size distribution depends, for given radiation quality \( Q \), target gas and pressure, and target geometry, only on the product \( d \rho \), which hereinafter is referred to as “distance” as well.

Ionisation cluster size distributions can be measured with nanodosimeters [2], such as the ion-counting nanodosimeter operated at PTB. The original setup of the device is described in detail in [3], and a detailed description of later improvements and an improved characterisation of the device are found in [4]. The nanodosimeter comprises an interaction region filled with a rarefied target gas, an electrode system to extract the ionised target gas molecules from the interaction region, an acceleration stage with an ion-counting detector at its end, and a primary particle detector. The interaction region is located between the electrodes of a plane parallel capacitor. A primary particle passing through the interaction region ionises the target molecules, and at the end of the interaction region it is registered in the primary particle detector, which starts the data acquisition. Due to an electric field across the plane parallel plate capacitor, the ionised target molecules are drifted towards the extraction electrode, which contains a small aperture. Ions created in a volume of cylindrical symmetry directly above this aperture are extracted from the interaction region and are accelerated through an electrode system towards an ion-counting secondary electron multiplier, where they are counted individually. Repeated measurements for a large number of primary particles yield...
the relative frequency distribution $P_\nu(Q, d\rho)$. In the following discussion, the relative frequency distribution of the target gas ions extracted through the extraction aperture is denoted $P^{\text{ION}}$ and the relative frequency distribution of the ions counted by the secondary ion multiplier is denoted $P^{\text{EXP}}$. For the sake of clarity, the dependency of the relative frequency distribution on radiation quality $Q$ and distance $d\rho$ is omitted in the following equations.

In previous investigations, a substantial deviation in the frequency of occurrence for large cluster sizes was found in the comparison between measured and simulated data of ionisation cluster size distributions [4, 5]. These deviations could be explained by a background consisting of additional ionisations, which are produced within the transport system for the ionised target gas molecules. Including this background of secondary ions in the simulated cluster size distributions led to a significantly better agreement between measured and simulated data, especially for large cluster sizes.

A correction of the simulation results with respect to the inclusion of the secondary ion production is sufficient for a mere comparison of the simulated with measured cluster size distributions. However, for the prediction of the biological effectiveness of different radiation qualities [6, 7], the “true” cluster size distributions must be determined. A procedure therefore needs to be applied to the measured data to remove the effects of the background due to secondary ionisations.

2 Modelling the background of secondary ions

The measured ionisation cluster size distributions contain a background which arises from secondary ions produced due to the scattering of extracted target gas ions within the ion transport optics downstream of the extraction aperture [4, 5]. In transport simulations, target gas ions were found to impinge on the surface of the electrodes of the ion transport system, having kinetic energies of up to 2.5 keV, after being scattered at the neutral target gas molecules expanding through the extraction aperture. Ions of this energy are able to create secondary electrons when impinging on a surface. Additionally, secondary electrons can be created in the collisions of the target gas ions with the neutral target gas molecules expanding through the extraction aperture. These electrons are accelerated towards the extraction aperture, gaining maximum energy in the region close to the extraction aperture where the gas density is maximal due to the target gas expanding out of the interaction chamber. These secondary electrons are able to create additional ions from the neutral target gas molecules expanding through the extraction aperture, which add to the ionised gas molecules originating from ionisations due to the primary particle.

Cosmic or terrestrial background radiation as well as spurious counts from the secondary electron multiplier can be excluded from being the source of the additional ions, as a corresponding investigation resulted in a total count rate of 0.06 per second. Instead, the background of additional ions is found to vanish almost completely, when the kinetic energy of the target gas ions inside critical regions of the ion transport optics, i.e. in regions with high target gas density or high kinetic energy of the target gas ions, is reduced substantially. This result clearly indicates a linkage between the kinetic energy of the target gas ions and the production of secondary electrons and secondary target gas ions, respectively, and supports the presented model of the secondary ion production. However, a detailed discussion of the investigation of the variation of the voltages applied to the different electrodes in the ion transport optics and its effect on the secondary ion background and,
moreover, its effect on the spatial distribution of the extraction efficiency, is beyond the scope of this manuscript and will be subject of another paper.

The ions originating from interactions of the primary particles in the nanodosimeter’s interaction volume typically arrive at the secondary electron multiplier within a time window of about 15 µs width at 10% of the peak frequency in the arrival time distribution. Secondary ions from ionisations by electrons, which were produced by ions that pass through the extraction aperture and then are scattered onto an electrode or which were produced in the collisions of the target gas ions with the neutral target gas molecules, are produced within a time span of about 2 µs after the initiating ion passed the extraction aperture. Hence, secondary ions cannot be distinguished from ions originating from primary particle interactions, and thus the size of the measured ionisation cluster increases.

The background in the experimental data is determined using a model [4] which is based on two quantities: the probability \( \varepsilon \) of an ionised target gas molecule to be scattered within the transport optics, and the expectation value \( \lambda \) of a Poisson distribution of the number of secondary ions created by the electrons released by a single ionised target gas molecule, which is scattered within the transport optics. Since the mechanism of secondary ion production is independent of the radiation quality of the primary particle, the two parameters \( \varepsilon \) and \( \lambda \) must have the same values for each cluster size distribution measured with any radiation quality for the specific measurement conditions, i.e. target gas, gas pressure and width of the drift time window. The “true” primary ionisation cluster size distributions created by the primary particles in the interaction volume do not include these secondary ions. For comparison with measured data, the “true” ionisation cluster size distributions need to be adjusted by including these secondary ionisations.

![Diagram](image)

**Figure 1.** Scheme of the inclusion of secondary ionisations in measured cluster size distributions.

The inclusion of secondary ionisations into the “true” primary ionisation cluster size distributions is a multi-step process (see figure 1). The first step is to split the primary cluster size distribution \( P^{ION} \) into two distributions according to the probability \( \varepsilon \) for an ionised gas molecule from an ionisation cluster to be scattered within the transport optics. The distribution \( P^{SEM} \) repre-
sents the ions which reach the secondary electron multiplier directly without being scattered. The other distribution $P^{SCT}$ represents those ions which are scattered within the transport optics. The latter distribution is, in the next step, combined with $P^{SEC}$ to obtain the background distribution $P^{BGD}$. $P^{SEC}$ is the set of distributions of the conditional probabilities that a certain number of secondary ions are produced given that a number $\nu$ of ions are scattered within the transport optics. For $\nu = 1$, $P^{SEC}$ is a Poisson distribution with an assumed expectation value $\lambda$ representing the number of secondary ions created by a single ionised target gas molecule, which is scattered within the transport optics. For $\nu > 1$, $P^{SEC}$ comprises all the $\nu$-fold convolutions of the aforementioned Poisson distribution with itself. $P^{SCT}$ and $P^{SEC}$ are combined such that the $\nu$-fold convoluted Poisson distributions are weighted according to the relative frequency of occurrence $P^{SCT}_{\nu}$ and are summed up. The result of this operation describes the background distribution $P^{BGD}_{\nu}$ of secondary ionisations created by electrons that are produced by primary target gas ions scattered within the transport optics. Convolution of $P^{BGD}_{\nu}$ and $P^{SEM}$ leads to the total distribution of ionisations $P^{ION+BGD}_{\nu}$, which can be compared to $P^{EXP}$ measured in the experiment.

Two approximations enter the model on which the construction of $P^{BGD}_{\nu}$ is based. The first is that the scattered target gas ions which cause the background of secondary ions are consumed in the process of releasing the secondary electrons which in turn create the secondary ions. Since the scattered target gas ions most likely are scattered out of the path leading to the secondary electron multiplier and are neutralised at the electrodes or the vacuum chamber walls, this assumption seems to be justified. The second approximation is that the secondary ions are detected in the secondary electron multiplier with the same efficiency as target gas ions passing through the extraction aperture. Due to the expansion of the target gas leaking out of the interaction chamber, the number density of target gas molecules decreases by orders of magnitude within a few millimetres from the extraction aperture. Hence, secondary ions are predominantly created in close vicinity to the extraction aperture where the energy of the secondary electrons is maximum. Therefore, the secondary ions are created at the beginning of the ion transport optics and can follow the path towards the secondary electron multiplier, where they arrive after acceleration to the full kinetic energy sufficient for detection with the same efficiency as the target gas ions passing through the extraction aperture.

The simple scheme shown in figure 1 is a simplification of the actual process. Since the secondary ions are created in the vicinity of the extraction aperture, one must take into account that these secondary ions can also be scattered by neutral gas atoms downstream of the extraction aperture and, thus, have a probability $\varepsilon$ of being scattered in the transport optics. Rather than being convoluted with $P^{SEM}$ to yield $P^{ION+BGD}$, the distribution $P^{BGD}$ has therefore to undergo the same process of splitting and convolution as described above in order to calculate the secondary ions of the next generation (see figure 2). This iterative procedure is continued until convergence is reached — that is, until the increase in the number of secondary ionisations due to an increasing number of generations becomes insignificant. Finally, the distributions from the splitting processes of all iterations representing the ions of the different generations reaching the secondary electron multiplier without scattering are convoluted with the total distribution of secondary ions of the last iteration step. At the conclusion of this process, the distribution $P^{ION+BGD}$ includes the contribution of all generations of secondary ions, which can subsequently be compared with measured data $P^{EXP}$.

For the approach within the framework of ORIGIN using the minimising NAG-routines.
Figure 2. Scheme of the iterative inclusion of consecutive generations of secondary ionisations in measured cluster size distributions.

“E04CCC” or “E04JBC”, the realisation of the model was quite straightforward: to apply “E04CCC” or “E04JBC”, the user has to provide an initial estimate of $P_{\text{ION}}$ and a routine to be called by the respective NAG-routine. The latter routine, which carries out the splittings and the convolutions according to the scheme in figure 2 to calculate the background-corrected estimate of the primary distribution, also contains the target function, i.e. the function describing the likelihood between the measured distribution $P_{\text{EXP}}$ and background-corrected estimate of $P_{\text{ION}}$, which is to be minimised.

The approach employing the iterative unfolding algorithm using Bayes statistics implies a constraint which complicates the mathematical formulation of the model: application of the algorithm described in [11, 12] requires a representation of the model in the form

$$P_{\text{ION} + \text{BGD}} = M \cdot P_{\text{ION}}$$

(2.1)

with the matrix $M$ describing the processes leading to the background of additional ions. The formulation in matrix notation of the different steps involved in the inclusion of secondary ionisations for the non-iterative, i.e. single generation, scheme shown in figure 1 are represented in the following set of equations.
Equations (2.2) and (2.3) describe the splitting process of the primary cluster size distribution $p_\text{ION}$ into the two distributions $p_\text{SCT}$ of ions scattered in the ion optics and $p_\text{SEM}$ of ions reaching the secondary electron multiplier without being scattered:

$$p_\text{SEM} = M_\text{SEM} \cdot p_\text{ION} \quad \text{with:} \quad M_\text{SEM}^{\nu_1, \nu_0} = \begin{cases} 0 & \text{if: } \nu_0 < \nu_1 \\ (1 - \varepsilon)^{\nu_1} \cdot e^{\nu_0 - \nu_1} \cdot \binom{\nu_0}{\nu_1} & \text{if: } \nu_0 \geq \nu_1 \end{cases}$$

$$p_\text{SCT} = M_\text{SCT} \cdot p_\text{ION} \quad \text{with:} \quad M_\text{SCT}^{\nu_2, \nu_0} = \begin{cases} 0 & \text{if: } \nu_0 < \nu_2 \\ (1 - \varepsilon)^{\nu_0 - \nu_2} \cdot e^{\nu_0 - \nu_2} \cdot \binom{\nu_0}{\nu_2} & \text{if: } \nu_0 \geq \nu_2 \end{cases}$$

The matrix elements $M_\text{SEM}^{\nu_1, \nu_0}$ are the conditional probabilities that $\nu_1$ ions reach the secondary electron multiplier without being scattered, provided that $\nu_0$ ions pass through the extraction aperture. Analogously, $M_\text{SCT}^{\nu_2, \nu_0}$ are the conditional probabilities that $\nu_2$ ions are scattered in the ion optics if $\nu_0$ ions pass through the extraction aperture. Since $\nu_0 = \nu_1 + \nu_2$, the two conditional probabilities are related by $M_\text{SEM}^{\nu_1, \nu_0} = M_\text{SCT}^{\nu_2, \nu_0}$.

The following equation (2.4) calculates the distribution $p_\text{BGD}$ of the background of secondary ionisations.

$$p_\text{BGD} = M_\text{SEC} \cdot p_\text{SCT} = M_\text{SEC} \cdot M_\text{SCT} \cdot p_\text{ION} \equiv M_\text{BGD} \cdot p_\text{ION}$$

with: $M_\text{SEC}^{\nu_1, \nu_2} = \frac{(\nu_2 - \lambda)^{\nu_1}}{\nu_1!} \cdot e^{-\nu_2 - \lambda}$

Equation (2.5) describes the convolution of $p_\text{SEM}$ and $p_\text{BGD}$ forming $p_\text{ION+BGD}$, which contains the background of secondary ionisations and is compared to the measured distribution $p_\text{EXP}$:

$$p_\text{ION+BGD} = M_\text{ADD} \cdot p_\text{SEM} = M_\text{ADD} \cdot M_\text{SEM} \cdot p_\text{ION} \equiv M \cdot p_\text{ION}$$

with: $M_\text{ADD}^{\nu_1, \nu_2} = \begin{cases} 0 & \text{if: } \nu_2 \geq \nu_1 \\ p_\text{BGD}^{\nu_2} & \text{if: } \nu_2 < \nu_1 \end{cases}$

The matrix product of $M_\text{ADD}$, which contains the contribution of $p_\text{BGD}$, and $M_\text{SEM}$ result in the matrix $M$ comprising all processes leading to the additional background of secondary ionisations.

In equation (2.5), the implicit assumption was made that all secondary ions would be registered at the secondary electron multiplier at the end of the detection system. However, as was pointed out above, the secondary ions produced in the vicinity of the extraction aperture may also be scattered in the ion optics when they collide with neutral gas molecules. This means that the procedure described above needs also be applied on the distribution $p_\text{BGD}$, resulting in a second generation of ions scattered in the ion optics. This will in turn lead to a second generation of secondary ions produced by electrons released from the electrodes or target gas molecules by ion impact, which are then accelerated to the vicinity of the extraction aperture. Taking into account the iterations, i.e. the
sequence of generations $N = 1 \ldots N_{\text{max}}$, in the production of secondary ionisations (see figure 2) results in another set of equations (with the elements of $M_{\text{SEM}}, M_{\text{SCT}}$ and $M_{\text{BGD}}$ being determined in the same way as previously). The meaning of these equations is similar to equations (2.2) through to (2.5), in which the splitting into sequences of generations of secondary ions was not accounted for.

Equations (2.6) and (2.7) describe the splitting process of the primary cluster size distribution $P_{\text{ION}}$ into the two distributions $P_{\text{SCT}}^N$ (ions scattered in the ion optics) and $P_{\text{SEM}}^N$ (ions reaching the secondary electron multiplier without scattering) of the $N^{th}$ generation, which take into account all previous generations:

$$P_{\text{SEM}}^N = M_{\text{SEM}} \cdot \prod_{n=1}^{N-1} M_{\text{BGD}} \cdot P_{\text{ION}}$$
$$P_{\text{SCT}}^N = M_{\text{SCT}} \cdot \prod_{n=1}^{N-1} M_{\text{BGD}} \cdot P_{\text{ION}}$$

The following equation (2.8) calculates the distribution $P_{\text{BGD}}^N$ of the background of secondary ionisations in the $N^{th}$ generation:

$$P_{\text{BGD}}^N = \prod_{n=1}^{N} M_{\text{BGD}} \cdot P_{\text{ION}}$$

This allows simplifying equations (2.6) and (2.7) to give:

$$P_{\text{SEM}}^N = M_{\text{SEM}} \cdot P_{\text{BGD}}^{N-1}$$
$$P_{\text{SCT}}^N = M_{\text{SCT}} \cdot P_{\text{BGD}}^{N-1}$$

The next equation (2.11) describes the convolutions of $P_{\text{SEM}}^N$ and $P_{\text{ADD}}^N$ of the contributions from the generations of secondary ions between the $N^{th}$ and the $(N_{\text{max}})^{th}$ generation in descending order, beginning with $N = N_{\text{max}}$ and ending with $N$:

$$P_{\text{ADD}}^{N-1} = M_{\text{ADD}}^N \cdot M_{\text{SEM}} \cdot \prod_{n=1}^{N-1} M_{\text{BGD}} \cdot P_{\text{ION}}^N = M_{\text{ADD}}^N \cdot M_{\text{SEM}} \cdot P_{\text{BGD}}^{N-1}$$

with:

$$M_{\text{ADD}}^N \left( x_1, x_2 \right) = \begin{cases} 0 & \text{if: } v_2 \geq v_1 \\ (P_{\text{ADD}}^N)_{(v_1, v_2)} & \text{if: } v_2 < v_1, N < N_{\text{max}} \\ (P_{\text{BGD}}^N)_{(v_1, v_2)} & \text{if: } v_2 < v_1, N = N_{\text{max}} \end{cases}$$

Finally, the distribution $P_{\text{ION+BGD}}$ (which technically corresponds to $P_{\text{ADD}}^0$) is obtained from $P_{\text{ION}}$ by multiplication with the matrix $M$ comprising all processes leading to the additional background of secondary ionisations and is compared to the measured distribution $P_{\text{EXP}}$:

$$P_{\text{ION+BGD}} = M_{\text{ADD}}^1 \cdot P_{\text{SEM}}^1 = M_{\text{ADD}}^1 \cdot M_{\text{SEM}} \cdot P_{\text{ION}} = M \cdot P_{\text{ION}}$$

The matrix product of $M_{\text{ADD}}$, which contains the contribution of all $P_{\text{BGD}}^N$, and $M_{\text{SEM}}$ results in the matrix $M$, which comprises all processes leading to the additional background of secondary ionisations.
Two target functions \( T \) were tested for use with “E04CCC” and “E04JBC”. They are defined by:

\[
T = \frac{1}{n} \sum_{\nu=0}^{\infty} \left[ \frac{P_{\nu}^{\text{EXP}} - P_{\nu}^{\text{ION+BGD}}}{P_{\nu}^{\text{EXP}} + P_{\nu}^{\text{ION+BGD}}} \right] \cdot W_{\nu}^{\text{EXP}}
\]  

(2.13)

where \( n \) is the number of cluster sizes with non-vanishing relative frequency. The two target functions differ by their definition of the weighting function \( W_{\nu}^{\text{EXP}} \): [4]

(i) \[
W_{\nu}^{\text{EXP}} = \begin{cases} 1 - 1/\sqrt{N_{\nu}^{\text{EXP}}} & \text{if } N_{\nu}^{\text{EXP}} > 0 \\ 0 & \text{if } N_{\nu}^{\text{EXP}} = 0 \end{cases}
\]  

(2.14)

(ii) \[
W_{\nu}^{\text{EXP}} = \frac{\sqrt{N_{\nu}^{\text{EXP}}}}{\sum_{\nu=0}^{\infty} N_{\nu}^{\text{EXP}}}
\]  

(2.15)

In the preceding equations, \( P_{\nu}^{\text{EXP}} \) is the measured relative frequency distribution of the ionisation cluster size and \( P_{\nu}^{\text{ION+BGD}} \) is the background-corrected estimate of \( P_{\nu}^{\text{ION}} \). The sum includes only those summands where both \( P_{\nu}^{\text{EXP}} > 0 \) and \( P_{\nu}^{\text{ION+BGD}} > 0 \), and \( n \) is the number of data points for which \( P_{\nu}^{\text{EXP}} > 0 \) and \( P_{\nu}^{\text{ION+BGD}} > 0 \). \( W_{\nu}^{\text{EXP}} \) describes the weight of summand \( \nu \), and \( N_{\nu}^{\text{EXP}} \) is the absolute frequency of measured ionisation clusters of size \( \nu \) which is related to the relative frequency by \( P_{\nu}^{\text{EXP}} = N_{\nu}^{\text{EXP}} / \sum_{\nu=0}^{\infty} N_{\nu}^{\text{EXP}} \).

The two weighting functions differ with respect to the range where they affect \( T \): the first weighting function is effective only in the range of very small \( N_{\nu}^{\text{EXP}} \), whereas the second one applies the classical \( \sqrt{N} \)-type weighting scheme affecting the whole range of \( N_{\nu}^{\text{EXP}} \).

Two different initial estimates \( P_{\nu}^{\text{ION},0} \) for \( P_{\nu}^{\text{ION}} \) were applied: (i) a uniform distribution of cluster sizes

\[
p_{\nu}^{\text{ION},0} = \frac{1}{m} \quad \text{if } 0 \leq \nu < m
\]  

(2.16)

and (ii) a distribution representing the mean value of the simulated probability \( P_{\nu}^{\text{SIM}} \) and measured probability \( P_{\nu}^{\text{EXP}} \) for each cluster size \( \nu \):

\[
p_{\nu}^{\text{ION},0} = \frac{P_{\nu}^{\text{EXP}} + P_{\nu}^{\text{SIM}}}{2}
\]  

(2.17)

with \( m \) in equation (2.16) representing the number of measured cluster sizes \( \nu \) with \( P_{\nu}^{\text{EXP}} > 0 \).

Due to the strict iteration scheme of the algorithm described in [11, 12], the user can not choose a target function \( T \) in the same sense as described when using the NAG-routines “E04CCC” or “E04JBC”. It is, however, possible to calculate the likelihood between the measured distribution \( P_{\nu}^{\text{EXP}} \) and background-corrected estimate \( P_{\nu}^{\text{ION+BGD}} \) according to a target function and subsequently select the iteration result that minimises the target function. As in the case of the NAG-routines, the user has to provide an initial estimate of \( P_{\nu}^{\text{ION}} \). For consistency, the same target functions \( T \) and the same initial estimates of \( P_{\nu}^{\text{ION}} \) were used as for “E04CCC” and “E04JBC”.

3 Results

In the present investigation, the above mentioned approaches to remove the effect of the background of secondary ionisations were applied to different sets of measurements with different values for \( \varepsilon \).
and \( \lambda \): (i) a set of measurements for a target gas of 1.2 mbar C\(_3\)H\(_8\) obtained for a number of radiation qualities (ion types and energies) with the primary particle hitting the target volume centrally and with a fully open drift time window \((\varepsilon = 0.0065, \lambda = 15)\) [4], (ii) a set for a target gas of 1.2 mbar N\(_2\) obtained with the primary particle hitting the target volume centrally and with a fully open drift time window \((\varepsilon = 0.045, \lambda = 3.2)\) [4] and (iii) a set of measurements for a target gas of 1.2 mbar C\(_3\)H\(_8\) obtained with the primary particle passing the target volume at distance \(d\rho\) and with a drift time window of \(\pm 2.5 \mu s\) \((\varepsilon = 0.014, \lambda = 9)\) [5]. For all data sets, the minimisation results were compared to the previously mentioned simulated cluster size distributions [4, 5], and the minimisation results were convoluted with the background of additional ions and then compared to the measured cluster size distributions.

The first part of the following discussion of the results focusses on both, the comparison of the minimisation results and the comparison of their convolution with the secondary ion background with the measured cluster size distributions, among the different methods. The second part deals with the effectiveness of the methods by comparing the minimisation results with the simulated data, and the third part focusses on the estimation of the uncertainties of the unfolded distributions obtained with the different methods.

### 3.1 Comparison of the unfolding methods

For “E04CCC”, the final results of the unfolding process strongly depend on the initial estimate of \(P^{\text{ION}}\) and only slightly on the weighting function \(W^{\text{EXP}}_\nu\) of the target function. For a uniformly-distributed initial estimate, no satisfactory agreement between the unfolding result convoluted with secondary ion background and the respective measured distribution could be obtained for any measurement, as shown in figure 3. On the other hand, for an initial estimate representing the mean of simulated and measured probability distribution, the majority (75%–80%) of the unfolding

![Figure 3](image-url)

*Figure 3.* Unsatisfactory agreement between the unfolding result obtained with “E04CCC” convoluted with secondary ion background and the measured cluster size distribution. Additional parameters: measurement: 420 keV protons in 1.2 mbar C\(_3\)H\(_8\), \(d\rho = 0 \mu g/cm^2\); unfolding: \(W^{\text{EXP}}_\nu\): (ii), \(P^{\text{ION},0}\): (i).
Table 1. Percentage of unfolding results convoluted with secondary ion background that agree well with the measurements for the different unfolding methods in relation to initial estimate $P_{\text{ION}}^0$ and weighting function $W_{\nu}^{\text{EXP}}$.

| $P_{\text{ION}}^0$ | $W_{\nu}^{\text{EXP}}$ | “E04CCC” | “E04JBC” | “BAYES” |
|---------------------|-------------------------|------------|-----------|----------|
| uniform (i)         | small $N_{\nu}^{\text{EXP}}$ (i) | 0%         | 20%       | 100%     |
| √$N$-type (ii)      | 0%                      | 40%        | 100%      |          |
| mean (ii)           | small $N_{\nu}^{\text{EXP}}$ (i) | 75%        | 80%       | 100%     |
| √$N$-type (ii)      | 80%                     | 100%       | 100%      |          |

results convoluted with secondary ion background agrees well with the measurements for both weighting functions (see table 1). Due to the difference in the weighting functions, the results of the unfolding differ slightly.

A similar behaviour is found for “E04JBC”. Even though the results of the unfolding still strongly depend on the initial estimate of $P_{\text{ION}}$, the dependency is much less pronounced. For a uniformly-distributed initial estimate, a good agreement with the measurement was achieved for 20% and 40% of the unfolding results convoluted with secondary ion background for both weighting functions. An initial estimate representing the mean of simulation and measurement led to a good agreement with the measurement for 80% and 100% of the unfolding results convoluted with secondary ion background. For both types of initial estimates, the larger number of unfolding results agreeing with measurements was obtained for the weighting function (ii), which applies the classical $\sqrt{N}$-type weighting scheme. As for “E04CCC”, the unfolding results differ slightly for the two weighting functions due to the difference in weighting.

The unfolding results convoluted with secondary ion background that did not agree with corresponding measurements were due to “E04JBC” and “E04JBC” converging into local minima. For a given initial estimate of $P_{\text{ION}}$ and given weighting function $W_{\nu}^{\text{EXP}}$, increasing the maximum number of minimising iterations did not free the routines out of the local minima where they were stuck. The only way to prevent the routines from converging into the respective local minima was to provide another initial estimate and/or weighting function.

For the iterative unfolding algorithm using Bayes statistics (denoted “BAYES” in the following text), a completely different behaviour was observed: the final results were almost independent from both the initial estimate of $P_{\text{ION}}$ and the weighting function $W_{\nu}^{\text{EXP}}$. Furthermore, all final results convoluted with secondary ion background of this algorithm agreed well with the corresponding measurement.

However, a peculiarity in the behaviour of the algorithm using Bayes statistics was observed. When secondary ionisations of the second generation or higher were included, the iteration was observed to become unstable, i.e. it diverged and/or oscillated. A stable operation could only be achieved when including only secondary ionisations of the first generation.

A possible explanation might lie in the dynamics of the matrix $M$, which comprises the processes leading to the additional background of secondary ionisations. When taking into account only secondary ions of the first generation, $M$ is a static matrix, i.e. its elements do not change between iteration steps. This is no longer the case when taking into account the secondary ions of higher generations. In this case, the inclusion of generations higher than the first leads to a matrix
Figure 4. Comparison of unfolding results obtained with the different methods convoluted with secondary ion background and the measured cluster size distribution (left) and comparison between the unfolding results (right). Additional parameters for figure 4a): measurement: 88 MeV C-ions in 1.2 mbar C$_3$H$_8$, $dp = \pm 1.71$ µg/cm$^2$; unfolding: $W^{\text{EXP}}$: (ii), $P^\text{ION,0}$: (ii). Additional parameters for figure 4b): measurement: 1.1 MeV alpha particles in 1.2 mbar N$_2$, $dp = 0$ µg/cm$^2$; unfolding: $W^{\text{EXP}}$: (ii), $P^\text{ION,0}$: (ii).

M with elements changing dynamically between iteration steps, and the dynamics increases with increasing generation depth.

The iteration formula (iteration $i$) of the algorithm using Bayes statistics [11, 12]:

$$P^\text{ION,(i+1)}_v = \frac{1}{\sum_{\mu} M_{\mu,v}} \cdot \sum_{\mu} M_{\mu,v} \cdot \sum_k M_{\mu,k} \cdot P^\text{EXP}_k \cdot P^\text{ION,i}_v$$ \hspace{1cm} (3.1)

can be interpreted as:

$$P^\text{ION,(i+1)}_v = \text{“step width”} \cdot P^\text{ION,i}_v$$ \hspace{1cm} (3.2)

with: \text{“step width”} = \frac{1}{\sum_{\mu} M_{\mu,v}} \cdot \sum_{\mu} \frac{M_{\mu,v} \cdot P^\text{EXP}_\mu}{\sum_k M_{\mu,k} \cdot P^\text{ION,i}_k}$$ \hspace{1cm} (3.3)
Empirically, a reduction of the step width by a factor taking into consideration the generation depth was found to restore the iteration’s stability. The new reduced step width was defined by:

\[
\text{“new step width”} = 1 + (\text{“step width”} - 1)/N_{\text{max}}
\]  

with \(N_{\text{max}}\) being the maximum generation depth.

Generally, when the unfolding results convoluted with the background of additional ions agree well with the corresponding measured cluster size distributions, only small or almost no deviations are found between the unfolding results for all methods employed. The left plots of figures 4a and 4b show two examples of a comparison between unfolding results of the three different methods convoluted with the background of secondary ions and measured distributions. The convoluted unfolding results of the different methods are almost indistinguishable from each other and agree very well with the measured cluster size distributions. A comparison of the unfolding results can be seen on the right. In figure 4a, the unfolding results are almost identical for the different methods, whereas in figure 4b differences can be seen for the largest cluster sizes in the distribution with the smallest frequency of occurrence, i.e. for cluster sizes \(\nu > 12\). Generally, if significant differences are observed for unfolding results whose convolutions with the secondary ion background agree well with the measurements, then these differences are found only in that region of the distribution with the largest cluster sizes having the smallest frequency of occurrence, as seen in figure 5.

In most cases, the unfolding results obtained with “BAYES” are independent of the initial estimate of \(P^\text{ION}\) and the weighting function, as shown in figure 6. Both the unfolding results and their convolution with the secondary ion background are identical. However, there are a few cases where a deviation between the results for the different initial estimates and weighting functions can be observed, as seen in figure 7. Here, the unfolding results differ slightly for the large cluster sizes at the end of the distribution, but the convolutions with the secondary ion background are almost identical.

![Figure 5](image)

**Figure 5.** Comparison of unfolding results obtained with the different methods convoluted with secondary ion background and the measured cluster size distribution (left) and comparison between the unfolding results (right). Additional parameters: measurement: 750 keV alpha particles in 1.2 mbar N\(_2\), \(d\rho = 0\) µg/cm\(^2\); unfolding: \(W^{\text{EXP}}_\nu\); (ii), \(P^\text{ION,0}\); (ii).
Figure 6. Comparison of unfolding results obtained with “BAYES” convoluted with secondary ion background and the measured cluster size distribution (left) and comparison between the unfolding results (right) for different \( W_{s \nu}^{\text{EXP}} \) and initial estimates \( P_{\text{ION}, 0} \). Additional parameters: measurement: 2.2 MeV alpha particles in 1.2 mbar \( C_3H_8 \), \( d \rho = 0 \, \mu g/cm^2 \).

Figure 7. Comparison of unfolding results obtained with “BAYES” convoluted with secondary ion background and the measured cluster size distribution (left) and comparison between the unfolding results (right) for different \( W_{s \nu}^{\text{EXP}} \) and initial estimates \( P_{\text{ION}, 0} \). Additional parameters: measurement: 750 keV alpha particles in 1.2 mbar \( N_2 \), \( d \rho = 0 \, \mu g/cm^2 \).

Figure 8 shows the unfolding results obtained with “E04JBC” for the two weighting functions using an initial estimate of \( P_{\text{ION}} \), which represents the mean between measured and simulated data. Both the unfolding results and their convolution with the secondary ion background appear to be almost identical. In cases where deviations between the measured distribution and the unfolding results convoluted with secondary ion background are observed, a similar trend can be observed (see figure 9). For the weighting function (i), the deviations between the unfolding result convoluted with the secondary ion background and the measurement are localised mainly in the range of cluster...
Figure 8. Comparison of unfolding results obtained with “E04JBC” convoluted with secondary ion background and the measured cluster size distribution (left) and comparison between the unfolding results (right) for different $W^\text{EXP}_\nu$ and initial estimates $P^\text{ION,0}$. Additional parameters: measurement: 250 keV protons in 1.2 mbar $\text{C}_3\text{H}_8$, $d\rho = 0 \mu\text{g/cm}^2$.

Figure 9. Comparison of unfolding results obtained with “E04JBC” convoluted with secondary ion background and the measured cluster size distribution (left) and comparison between the unfolding results (right) for different $W^\text{EXP}_\nu$ and initial estimates $P^\text{ION,0}$. Additional parameters: measurement: 88 MeV C-ions in 1.2 mbar $\text{C}_3\text{H}_8$, $d\rho = 0.34 \mu\text{g/cm}^2$.

sizes $\nu \leq 6$, with the unfolding result showing lower frequencies for cluster sizes $2 \leq \nu \leq 6$ and a larger frequency for $\nu = 0$. Additionally, in the range around $\nu \approx 15$ the unfolding result shows slightly larger frequencies than the measurement. The unfolding result convoluted with secondary ion background obtained with “E04JBC” for weighting function (ii), on the other hand, coincides with the measurement for cluster sizes $\nu \leq 11$. The deviations between measurement and unfolding result convoluted with secondary ion background are limited to cluster sizes $12 \leq \nu \leq 22$ with the
unfolding result showing larger frequencies than the measurement. This behaviour is reflected in the comparisons of the two unfolding results. Despite the unfolding result for weighting function (ii) truncating at smaller cluster sizes than the unfolding result for weighting function (i), the unfolding result for the former weighting function shows lower frequencies than the latter for cluster sizes \(2 \leq \nu \leq 6\) and a larger frequency for \(\nu = 0\). The reason for these differences in the unfolding results for the two different weighting functions can be attributed to their weighting schemes. Weighting function (i) reduces the weight only in the range of cluster sizes having very small frequencies, thus giving an almost identical weight to larger frequencies, whereas weighting function (ii) applies the classical \(\sqrt{N}\)-type weighting scheme, which emphasises those cluster sizes with larger frequencies the more the larger their frequency. A similar behaviour with respect to the two weighting functions is observed for the unfolding results obtained by “E04CCC”.

3.2 Comparison of the unfolding results with the simulated data

The degree of agreement between the simulated cluster size distributions and the results of the unfolding methods, which are applied with the aim to remove the background of secondary ions from the measured cluster size distributions, does not only depend on the method, the initial estimate of \(P_{\text{ION}}\) and the weighting function, but also, in particular, on the correctness and completeness of the model describing the development of the background of the additional ions. Due to the simplicity of the present model describing the background of secondary ions using only two parameters, which, moreover, are independent of the radiation quality but rather have the same value for all cluster size distributions measured for the same specific conditions (i.e. target gas, gas pressure and drift time window width), it cannot be expected that the background-corrected measurement, i.e. the result of the unfolding, and the corresponding simulation agree equally well in all cases.

Figure 10 shows the effect of the background correction, i.e. the removal of the background of additional ions, on the measured cluster size distributions exemplarily for some selected measurements. The simulated data are plotted in comparison with the results of the different unfolding methods, which represent the background-corrected measured data together with the uncorrected measured data. For the sake of clarity, the convolution of the unfolding results with the secondary ion background is omitted in this and subsequent figures. Nevertheless, the convolutions with the secondary ion background agree well with the measurements for all measured data presented.

In all cases shown in figure 10, the effect of the background correction on the measurements is clearly visible. For large cluster sizes, the background-corrected measured cluster size distributions show significantly lower frequencies than the not background-corrected distributions. Despite some scatter in the background-corrected data is observed for the different methods at cluster sizes with low frequencies, the background-corrected measurements are not only similar, or in some cases almost identical, for cluster sizes with high frequencies, but they are also close to the simulated distributions.

For large cluster sizes with low frequencies, the background-corrected cluster size distributions of a number of measurements differ from the simulated distributions. Here, two alternative findings are observed: either the frequencies in the background-corrected cluster size distributions are higher than in the simulated distributions, which can lead to a background-corrected distribution being in better agreement with the measured than the simulated distribution (figure 11), or the background-corrected distribution shows lower frequencies than those simulated (figure 12a) re-
Figure 10. Comparison of unfolding results obtained with the different methods with the simulated cluster size distribution and the measured distribution. All unfolding: $W^{\text{EXP}}_{\nu}$, PION$^0$, (ii). Additional parameters for figure 10a): measurement: 250 keV protons in 1.2 mbar C$_3$H$_8$, $d\rho = 0 \mu g/cm^2$. Additional parameters for figure 10b): measurement: 88 MeV C-ions in 1.2 mbar C$_3$H$_8$, $d\rho = 0.34 \mu g/cm^2$. Additional parameters for figure 10c): measurement: 88 MeV C-ions in 1.2 mbar C$_3$H$_8$, $d\rho = 0.85 \mu g/cm^2$. Additional parameters for figure 10d): measurement: 750 keV alpha particles in 1.2 mbar N$_2$, $d\rho = 0 \mu g/cm^2$.

sulting in truncated background-corrected distributions (figure 12b). These effects can be attributed to shortcomings of the model describing the background of additional ions due to its simplicity as discussed above.

A unique example is shown in figure 13. The background-corrected measured frequencies at cluster sizes larger than the cluster size of the maximum frequency, i.e. at cluster sizes $\nu \gtrsim 20$, agree well with the simulated data and show significantly smaller frequencies than the not background-corrected measurement. However, for cluster sizes $\nu \lesssim 20$, the background-corrected distribution differs significantly from the simulated but coincides with the not background-corrected measured distribution. The reason for this behaviour is that the parameter $\lambda$ used in the model for the
Figure 11. Comparison of unfolding results obtained with the different methods with the simulated cluster size distribution and the measured distribution. Additional parameters: unfolding: $W_{\nu}^{\text{EXP}}$: (ii), $P_{\text{ION}}^{0}$: (ii); measurement: 150 MeV C-ions in 1.2 mbar C$_3$H$_8$, $d\rho = \pm 1.71 \, \mu g/cm^2$.

Figure 12. Comparison of unfolding results obtained with the different methods with the simulated cluster size distribution and the measured distribution. All unfolding: $W_{\nu}^{\text{EXP}}$: (ii), $P_{\text{ION}}^{0}$: (ii). Additional parameters for figure 12a): measurement: 1.1 MeV alpha particles in 1.2 mbar N$_2$, $d\rho = 0 \, \mu g/cm^2$. Additional parameters for figure 12b): measurement: 3.2 MeV alpha particles in 1.2 mbar N$_2$, $d\rho = 0 \, \mu g/cm^2$.

description of the secondary ion background (mean number of secondary ionisations per target gas ion scattered in the ion optics) was always larger than one. Consequently, there will be more additional ions produced than are lost due to scattering processes, which shifts the frequency distribution towards larger cluster sizes.
Figure 13. Comparison of unfolding results obtained with “BAYES” for different $W_{\nu}^{\text{EXP}}$ and initial estimates $P^{\text{ION},0}$ with the simulated cluster size distribution and the measured distribution. Additional parameters: measurement: 440 keV alpha particles in 1.2 mbar $C_3H_8$, $d\rho = 0 \mu g/cm^2$.

3.3 Estimation of uncertainties

As none of the unfolding methods employed to remove the background of secondary ionisations from the measured data explicitly states uncertainties for the unfolding results, the uncertainties were estimated by comparison of the unfolding results obtained by the three methods for different initial estimates of $P^{\text{ION}}$ and weighting functions using a selection of previously shown unfolding results.

Figure 14a and figure 14b show the comparison of the unfolding results shown previously in figure 4a and figure 5, respectively. The plots show the mean of the unfolding results (black dots) together with the statistical uncertainties of the three methods on a logarithmic scale (right scale) and the ratio of the relative frequencies of the unfolding results of the respective method to the mean of the three unfolding results (coloured dots) on a linear scale (left scale). In figure 14a, the deviation of the relative frequency of the unfolding results of the different methods with respect to the mean is in the order of a few percent (except for cluster sizes $15 \leq \nu \leq 20$, where the relative deviation ranges between 5% and 12%) for cluster sizes of relative frequency larger than $10^{-4}$. In figure 14b, this deviation is in the order of a few percent for cluster sizes of relative frequency larger than $3 \cdot 10^{-3}$. Larger deviations are observed only for cluster sizes with much smaller frequency of occurrence.

To illustrate the effect of these deviations on quantities derived from the unfolded distributions, the first two moments were chosen. The $i$th moment of the distribution is defined by:

$$M_i = \sum_{\nu=0}^{\infty} \nu^i \cdot P_{\nu}. \quad (3.5)$$

For the unfolded distributions shown in figure 14a, $m_1$ varies between 1.3092 and 1.3294 and $m_2$ between 11.878 and 12.183, i.e. less than 1.6% and 2.6%, respectively. Similarly, for the distribution in figure 14b, $m_1$ varies between 4.2544 and 4.2568 and $m_2$ between 24.547 and 24.582, i.e. less than 0.06% and 0.15%, respectively.
Figure 14. Comparison of unfolded distributions obtained with the different methods with the corresponding mean distribution. Additional parameters for figure 14a): measurement: 88 MeV C-ions in 1.2 mbar C$_3$H$_8$, $d\rho = \pm 1.71$ µg/cm$^2$; unfolding: $W^{\text{EXP}}_\nu$, (ii), $P^{\text{ION},0}$. (ii). Additional parameters for figure 14b): measurement: 750 keV alpha particles in 1.2 mbar N$_2$, $d\rho = 0$ µg/cm$^2$; unfolding: $W^{\text{EXP}}_\nu$, (ii), $P^{\text{ION},0}$. (ii).

Figure 15a and figure 15b show the comparison of the unfolding results shown previously in figure 7 and figure 8, respectively. The plots show the deviations of the unfolded distribution for different initial estimates of $P^{\text{ION}}$ and the weighting functions $W^{\text{EXP}}_\nu$ for “BAYES” (figure 15a) and “E04JBC” (figure 15b) from the corresponding mean of the unfolded distributions. The unfolding results obtained with “BAYES” are almost identical for cluster sizes $\nu \leq 10$. Larger deviations from the mean distribution are only observed for cluster sizes of relative frequency less than $4 \cdot 10^{-4}$. For the unfolding results obtained with “E04JBC” the deviations are larger than for “BAYES”. However, for cluster sizes with frequencies larger than $3 \cdot 10^{-3}$, the deviations are less than 10%. Larger deviations are only observed for cluster sizes having smaller frequencies of occurrence.

Figure 15. Comparison of unfolded distributions obtained for different $W^{\text{EXP}}_\nu$ and initial estimates $P^{\text{ION},0}$ with the corresponding mean distribution. Additional parameters for figure 15a): unfolding: “BAYES”; measurement: 750 keV alpha particles in 1.2 mbar N$_2$, $d\rho = 0$ µg/cm$^2$. Additional parameters for figure 15b): unfolding: “E04JBC”; measurement: 250 keV protons in 1.2 mbar C$_3$H$_8$, $d\rho = 0$ µg/cm$^2$. 
The effect on the moments of the unfolded distributions are minor: for the unfolded distributions obtained by “BAYES”, $m_1$ varies between 4.25658 and 4.25682 and $m_2$ between 24.5786 and 24.5822, i.e. less than 0.006% and 0.015%, respectively. Similarly, for the unfolded distribution obtained with “E04JBC”, $m_1$ varies between 8.528 and 8.644 and $m_2$ between 94.95 and 96.76, i.e. less than 1.4% and 1.9%, respectively.

An extreme example for the deviation between the unfolded distributions obtained with the three methods and the corresponding mean distribution is shown in figure 16, which uses the data previously shown in figure 10b. Here large deviations between the coinciding unfolded distributions obtained by “BAYES” and “E04JBC” on the one hand and the distribution obtained with “E04CCC” on the other hand are observed for the complete range of cluster sizes, leading to $m_1$ varying between 1.36 and 1.75 and $m_2$ between 4.81 and 6.76, i.e. by about 30% and 40%, respectively. Considering that “E04CCC” is the routine which is least robust and most sensitive to the choice of the initial estimate of $P_{ION}$ and weighting function, the comparison shown in figure 16 is not representative.

The comparison of the unfolded distributions obtained with the different methods shows that deviations larger than 10% of the unfolded distributions from the corresponding mean distribution in general are limited to cluster sizes having a small frequency of occurrence. Since the amount with which the deviations contribute to the unfolded distributions is proportional to the frequency of occurrence of the respective ionisation clusters, the effect of the deviations on quantities derived from the unfolded distributions, e.g. the moments of the distributions, is generally in the order of a few percent or less.

The above comparison allows to estimate only uncertainties of the unfolded distributions due to the unfolding method, the initial estimate of $P_{ION}$ and the weighting function. It does not, however, allow conclusions with respect to the suitability of the underlying model describing the background of secondary ionisations. Nevertheless, in view of the comparison of unfolded distributions obtained with “E04JBC” and “BAYES” with simulated distributions, the comparatively simple model describing the background of secondary ions seems feasible to correct measured cluster size distributions with respect to the background of secondary ions.
4 Conclusion

Two different approaches were investigated to background-correct measured cluster size distributions with respect to a background of secondary ions. In the first approach, the removal of the background was treated as a minimising problem. This approach was carried out within the framework of the commercial software package ORIGIN using the minimising NAG-routines “E04CCC” and “E04JBC”. In the second approach, an iterative unfolding algorithm using Bayes statistics was employed. These approaches to background-correct measured cluster size distributions were applied to different sets of measurements with different values for $\varepsilon$ and $\lambda$.

For all methods and cases investigated, the background-correction of the measured cluster size distributions led to an improved agreement between background-corrected measurement and simulation as compared to the agreement between not background-corrected measurement and simulation, providing the background correction could be carried out successfully, i.e. the convolution of the unfolding results with the secondary ion background agreed well with the corresponding measured cluster size distribution. The degree of improvement differed for different cluster size distributions. In most cases, the background was successfully corrected and the background-corrected measured distribution agreed well with the simulated distributions. This was also observed for cases where the background-corrected measurement differed from the simulated distribution as these differences, which are due to shortcomings of the underlying simple model of the secondary ion background, mostly affect large cluster sizes with low frequencies. For only a few measurements did the background correction lead to a difference between background-corrected measurement and simulation for cluster sizes less than those occurring with maximum frequency. The reason for these differences is that the parameter $\lambda$ used in the model describing the mean number of secondary ions produced per target gas ion scattered in the ion optics was always larger than one, so that there were more additional ions produced than are lost due to scattering processes. This results in shifting the frequency distribution towards larger cluster sizes.

The different approaches differ with respect to the amount of successfully processed background corrections. The unfolding algorithm using Bayes statistics proved to be the most effective and the least sensitive to a weighting function and initial estimate of $P^{\text{ION}}$. All final results convoluted with their respective secondary ion background agree well with the corresponding measurement and they are almost independent from both the initial estimate of $P^{\text{ION}}$ and the weighting function. However, to operate this algorithm successfully and restore a stable iteration, the step width of the iteration had to be modified depending on the number of generations of secondary ions included.

In terms of effectiveness, the second-best correction was achieved with “E04JBC”. A uniformly-distributed initial estimate led to a good agreement with the measurement for 20% and 40% of the unfolding results convoluted with secondary ion background for both weighting functions. An initial estimate representing the mean of simulation and measurement also led to a good agreement with the measurement for 80% and 100% of the unfolding results convoluted with secondary ion background. The larger numbers were obtained for the weighting function (ii), which applies the classical $\sqrt{N}$-type weighting scheme.

The least effective and the most sensitive to initial estimate of $P^{\text{ION}}$ was “E04CCC”. For a uniformly-distributed initial estimate no satisfactory agreement between unfolding result convoluted with secondary ion background and measurement could be obtained for any measurement. An
initial estimate representing the mean of simulated and measured probability distribution led to an agreement with the measurement for 75%–80% of the unfolding results convoluted with secondary ion background for both weighting functions.

Comparison of the unfolded distributions obtained with the different methods shows that large deviations of the unfolded distributions from the corresponding mean distribution are generally limited to large clusters having a small frequency of occurrence. Since the contributions of the deviations to the unfolded distributions is proportional to the frequency of occurrence of the respective ionisation clusters, the effect of the deviations on quantities derived from the unfolded distributions generally is in the order of less than a few percent.

In view of the results of “E04JBC” and, in particular the unfolding algorithm using Bayes statistics, the correction of measured cluster size distributions with respect to the background of secondary ions created in the ion transport optics of the nanodosimeter seems feasible, despite the comparatively simple model describing the background of secondary ions. The background-corrected cluster size distributions can be applied in a situation when either the “true” measured cluster size distribution or derived quantities obtained from measurements are required, e.g. for the prediction of the biological effectiveness of different radiation qualities.

For this unique problem of background-correction of measured cluster size distributions, the unfolding algorithm using Bayes statistics is the preferred method as it proved to be the most effective and the least sensitive to weighting function and initial estimate of $P_{\text{ION}}$. Moreover, it was by far the quickest.

Acknowledgments

The authors gratefully acknowledge the developers of the device from the Weizmann Institute of Science, Rehovot, Israel, for transferring the nanodosimeter as described in [3] to PTB for further use. The authors would like to thank the staff of the ion accelerator facilities at the LNL-INFN and at PTB for their assistance and support during the measurements, and Dr. D. Moro (formerly at LNL-INFN) for his assistance and support in the preparation of the measurements. The authors also express their gratitude to W. Helms†, B. Lambertsen and A. Pausewang for their invaluable contributions in preparing and carrying out of measurements, as well as their assistance in data processing. H. Nettelbeck’s proofreading of the manuscript is also gratefully acknowledged.

This work was partly funded within EMRP Joint Research Project SIB06 BioQuaRT and EC grant agreement 262010 (ENSAR). The EMRP was jointly funded by the EMRP participating countries within EURAMET and the European Union.

References

[1] R. Schulte, Nanodosimetry: Principle and Current Status, AIP Conf. Proc. 1345 (2011) 249.
[2] A. Bantsar, P. Colautti, V. Conte, G. Hilgers, M. Pietrzak, S. Pszona et al., State of the Art of Instrumentation in Experimental Nanodosimetry, Radiat. Prot. Dosim. 180 (2018) 177.
[3] G. Garty, S. Shchemelinin, A. Breskin, R. Chechik, G. Assaf, I. Orion et al., The performance of a novel ion-counting nanodosimeter, Nucl. Instrum. Meth. A 492 (2002) 212.
[4] G. Hilgers, M.U. Bug and H. Rabus, Secondary ionizations in a wall-less ion-counting nanodosimeter: quantitative analysis and the effect on the comparison of measured and simulated track structure parameters in nanometric volumes, *Eur. Phys. J. D* **69** (2015) 239.

[5] G. Hilgers, M.U. Bug and H. Rabus, Measurement of track structure parameters of low and medium energy helium and carbon ions in nanometric volumes, *Phys. Med. Biol.* **62** (2017) 7569.

[6] V. Conte, A. Selva, P. Colautti, G. Hilgers and H. Rabus, Track structure characterization and its link to radiobiology, *Radiat. Meas.* **106** (2017) 506.

[7] V. Conte, A. Selva, P. Colautti, G. Hilgers and H. Rabus, Nanodosimetry: Towards a New Concept of Radiation Quality, *Radiat. Prot. Dosim.* **180** (2018) 150.

[8] http://www.originlab.com/.

[9] https://www.originlab.com/pdfs/nagcl09/manual/pdf/E04/e04ccc.pdf.

[10] https://www.originlab.com/pdfs/nagcl25/manual/pdf/e04/e04jbc.pdf.

[11] T.J. Kennett, W.V. Prestwich and A. Robertson, Bayesian deconvolution i: Convergent properties, *Nucl. Instrum. Meth.* **151** (1978) 285.

[12] G. D’Agostini, A multidimensional unfolding method based on Bayes’ theorem, *Nucl. Instrum. Meth. A* **362** (1995) 487.