JAAS

PAPER

View Article Online
View Journal | View Issue

Cite this: *J. Anal. At. Spectrom.*, 2023, **38**, 2607

Improving detection thresholds and robust event filtering in single-particle and single-cell ICP-MS analysis†

Matthias Elinkmann, *[a] Sarah Reuter,[b] Michael Holtkamp,[a] Steffen Heuckeroth, [a] Alexander Köhrer, [a] Katharina Kronenberg, [a] Michael Sperling, [c] Oliver Rubner, [d] C. Derrick Quarles, Jr, [e] Michael Hippler[b] and Uwe Karst [a]

In this work, a modular data processing workflow for single-particle (sp) and single-cell (sc) inductively coupled plasma-mass spectrometry (ICP-MS) is presented. To achieve more reliable detection thresholds, a special focus is placed on the parameter estimation of Gaussian and Poisson distributions that describe the background (BG) signal. For Gaussian models, the widely used iterative outlier test was improved by an algorithm that adjusts the test for different BG levels by incrementing the test factor. Through careful evaluation, the standard deviation of the experimental sc- and sp-ICP-MS data was applied as a robust measure of the convergence quality of the test. In addition, the outlier analysis was separated from the subsequent event detection more strictly than it is often reported. Importantly, a data-dependent decision criterion based on Gaussian and Poisson modeling was developed to effectively address extra-Poisson variance in experimental data. In the second part, a gate filter was developed to reduce the amount of excess false-positive events in sc-ICP-MS. To that end, a secondary filter based on the signal peak height is used to remove rare false-positive events without affecting the signal intensity of the events that are detected correctly. Two approaches, based on a numerical approximation *via* the detection limit, and critical values of the Gaussian and Poisson distribution are presented to calculate the gate filter level. Possible sources of false-positive events, some of which are specific to sc-ICP-MS, are discussed. The combined processing workflow was applied to analyze the distribution of six endogenous elements in *Chlamydomonas reinhardtii* cell populations. The gate filter corrected the cell number concentration by up to 44% (22% on average), and mass per cell by up to 30% (17% on average).

Received 29th August 2023
Accepted 25th October 2023

DOI: 10.1039/d3ja00292f

rsc.li/jaas

Introduction

Single-particle and single-cell ICP-MS

Inductively coupled plasma-mass spectrometry (ICP-MS) has been established as a versatile analytical technique for the elemental analysis of nanomaterials on a single particle (sp) level. In recent publications, it has been applied to answer complex ecological and toxicological questions with nanoparticles (NPs) of mixed composition or high polydispersity, and various surface modifications.[1–5] A promising variant of sp-ICP-MS is the analysis of biological cells, which enables the detection of chemical elements at the single-cell (sc) level. It is referred to as sc-ICP-MS. Due to challenges in cell handling, a lot of progress has been achieved through the optimization of sample introduction and preparation.[6,7] Here, the green algal strain *Chlamydomonas reinhardtii* (*C. reinhardtii*) proved to be a reliable test organism.[8,9] At the same time, data processing and, in particular, the definition of a suitable detection threshold remain a challenge with great research interest for both methods.[10–13] In this work, a modular data processing strategy for sp- and sc-ICP-MS is presented. The fundamentals and applicability of key statistical concepts are discussed, with emphasis on the specific requirements of sp- and sc-ICP-MS. The overall goal is to outline an algorithm that reliably characterizes the distribution of the background signal, detects cell events, and improves the determination of the element mass distribution and cell concentration.

[a]*Institute of Inorganic and Analytical Chemistry, University of Münster, Corrensstr. 48, 48149 Münster, Germany. E-mail: m.elin@uni-muenster.de*

[b]*Institute of Plant Biology and Biotechnology, University of Münster, Schlossplatz 8, 48143 Münster, Germany*

[c]*European Virtual Institute for Speciation Analysis (EVISA), Corrensstr. 48, 48149 Münster, Germany*

[d]*Institute of Physical Chemistry, University of Münster, Corrensstr. 28/30, 48149 Münster, Germany*

[e]*Elemental Scientific, Inc., 7277 World Communications Dr, Omaha, NE 68022, USA*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3ja00292f

Detection criteria

In atomic spectroscopy, Gaussian and Poisson models are widely used to decide if a signal is greater than the background and can be considered detected. Gaussian refers to a normal distribution with the parameters μ for the location and σ for the spread. The Poisson distribution has only one parameter μ and describes the probability of observing a positive integer number of events (0, 1, 2, …) in a given interval. It is a powerful model for memoryless counting processes such as the ion detection in an electron multiplier. In contrast to Gaussians, the Poisson distribution excludes negative values and has a pronounced positive skew (asymmetry) when μ gets small.[14–17] In both models, μ is usually estimated with the sample mean of experimental data, and the Gaussian σ with the empirical sample standard deviation (SD). For $\mu \geq 10$ however, the Poisson distribution can be approximated with a Gaussian whose variance (σ^2) is set equal to the mean, ($\sigma^2 = \mu$). This gets less accurate for smaller μ but has been reported for $\mu \geq 5$ and with some precaution even lower than 5.[13,17]

Kaiser popularized the widely known "3σ" criterion as a limit of detection based on Gaussian significance levels. Initially developed for optical spectroscopy, the concept was generalized to evaluate a "complete analytical method" based on the standard deviation of a series of blank samples and also for time-resolved signal acquisition.[14,15] Around the same time, Currie presented a rigorous formalism of detection criteria with a focus on non-continuous detection mechanisms like radioactive decay counting.[16] Accordingly, one-sided significance levels define a critical limit (L_C) based on the α-error (α: false-positive proportion) while the detection limit (L_D) factors in both α- and β-error (β: false negative proportion). This definition allows that the upper α% of background signals would be falsely identified as true signals, and the lower β% of true signals would be misinterpreted as background (BG). Formally, L_C and L_D refer to the net signal after a blank correction, and the corresponding expressions for the gross signal are Y_C and Y_D. Panel A in Fig. 7 contains a graphic representation of this concept, and Table 1 summarizes the expressions given by Currie for the "well-known" case, which will be discussed in the following section.[16] In addition, the Gaussian model assumes equal variances of true signal and BG, and the Poisson model uses the approximation $\sigma^2 = \mu$. One-sided z-scores for the standard normal distribution, that are a measure of the area under the Gaussian function, are applied for both models with $z_{1-\alpha} = z_\beta$. Interestingly, the Gaussian L_D obtained as "3.29σ" is numerically close to Kaiser's "3σ". The corresponding Poisson expression differs by a constant summand of $z^2 = 2.71$, resulting from the heteroskedasticity of the distribution.[16] These formulae are useful because the exact Poisson limits are not common in all software packages, and there is no analytic expression for L_D.[11,16]

In sp- and sc-ICP-MS, detection thresholds decide if the pulse height of a time-resolved signal is large enough to be counted as part of an event. The parameters of the models are often estimated with the mean (μ_B) and SD (σ_B) directly from the time-resolved BG signal. As an example, one of the first publications on liquid sample introduction sp-ICP-MS by Degueldre *et al.* reports the "3σ_B" criterion as the particle event detection threshold, which has been widely adopted.[18–21] As an alternative, the Poisson expressions in Table 1 were applied to take advantage of the ion counting-based signal from the secondary electron multiplier detector in ICP-MS.[12,22,23] More recent instruments are particularly suitable for Poisson statistics because shorter dwell times (DTs) are possible, generating only a few counts per reading.[11,12,23] Usually, the time-resolved signal is not subjected to a net correction, and the gross expressions (Y_C, Y_D) apply. Alternatively, empirical detection thresholds have been reported for NP samples based on the quantiles of a blank sample signal.[24] In sc-ICP-MS, however, cell suspensions can hardly be simulated with a blank due to deterioration processes. Since a Gaussian or Poisson model predicts probabilities without requiring an observation in a blank sample, this approach was preferred in this work.

Outlier tests

In the context of sp- and sc-ICP-MS, each data point of an event signal represents an outlier that distorts the estimated μ_B and σ_B of the BG distribution. For Poisson models, the mean may be replaced by the median which is more robust against outliers.[12] At low signal levels, however, the median is trivially always equal to zero. For Gaussian models, which require σ, the problem is predominantly solved by iteratively removing the outliers based on the "3σ_B", or more generally, "$f \times \sigma_B$" criterion.[19,25,26] Here, f represents a constant factor. This method calculates the mean and SD of the intensity data, removes all data points outside a range of "$\mu_B \pm f \times \sigma_B$" and repeats this procedure until no points outside the range remain. In the literature, a wide range of values for f from 1 to 10 is reported, and the same f is usually applied for both the outlier test and the detection threshold.[18,25,27–29] Tuoriniemi *et al.* investigated optimization strategies that systematically evaluate the same data set for different integer values up to $f = 20$, choosing $f = 5$ to minimize false-positive events for most analyses.[30,31] Hendriks *et al.* pointed out that a factor of $f > 5$ may indicate a mismatch

Table 1 Equations for the "well-known" case L_C and L_D for $\alpha = \beta = 5\%$.[16] The general definition uses z-scores for the standard normal distribution. In the specific expressions, μ_B and σ_B, denote the mean and sample standard deviation of the background signal, respectively. L_C and L_D are net value expression. The corresponding gross expressions are Y_C and Y_D

Definition	Gross expression		Gaussian model		Poisson model	
$L_C = z_{1-\alpha} \times \sigma$	$Y_C = L_C + \mu_B$	(1)	$L_C = 1.645 \times \sigma_B$	(2)	$L_C = 1.645 \times \sqrt{\mu_B}$	(3)
$L_D = L_C + z_\beta \times \sigma$	$Y_D = L_D + \mu_B$	(4)	$L_D = 3.29 \times \sigma_B$	(5)	$L_D = 2.71 + 3.29 \times \sqrt{\mu_B}$	(6)

between the data and a Gaussian model because such an event would be simply too rare.[32] In this work, the iterative approach is examined in detail to identify failure points and develop an improved procedure.

The "well-known" background

In the following section, Currie's statistical approach is reviewed with respect to the detection criteria in sp- and sc-ICP-MS. According to the definition, the null hypothesis H_0 states that the difference of the observed detector response (I_{S+B}) and the underlying blank signal (I_B) is zero. In that case, no net signal (I_S) would be present.[16] Such a difference of two random variables ($d = I_{S+B} - I_B$) can be evaluated with a paired difference test, hence the term "paired observations" given in the literature.[11,16] The reasoning is analogous to comparing a pair of sample and corresponding blank as discussed by Kaiser.[14] Using propagation of uncertainty, the variance of the difference can be estimated as ${\sigma_d}^2 = {\sigma_{S+B}}^2 + {\sigma_B}^2$. Experimentally, σ_B is readily available from observations of the BG signal or suitable blank samples. Assuming similar variances for signal levels of I_B and I_{S+B}, the variance of the observed detector response σ_{S+B} can also be approximated by σ_B, resulting in $\sigma_d = \sqrt{2}\cdot\sigma_B$.[16] Alternatively, the BG can be characterized through a long series of observations. Such a "well-known" BG provides reliable estimates of μ_B and σ_B. Instead of a pairwise comparison, the "well-known" BG is treated as a constant signal offset. For the net signal difference, the BG does not represent a random variable ($d = I_{S+B} - \mu_B$), which reduces the variance estimate to $\sigma_d = \sigma_{S+B}$. Finally, the "well-known" σ_B is again used as the estimator for σ_{S+B}, giving $\sigma_d = \sigma_B$.[16] eqn (2) could hence be interpreted as $L_C = 1.645 \times \sigma_d$, corresponding to the 95th percentile of cases in that no net signal would be present. In sp- and sc-ICP-MS, the associated events are rare and short intensity spikes. Consequently, most data points are pure background signal. By removing the outliers, the large number of remaining data points can be reasonably considered a "well-known" background. For instance, with a DT of 1 ms, a sampling time of 60 s generates 6×10^4 observations. With a typical event rate of $\leq 50\ s^{-1}$ and event duration of 0.5 ms, there are only $\leq 4.5 \times 10^3$ detector readings associated with an event.[33,34] Following a similar argument, test statistics for small sample sizes like a *t*-test are not required because the large number of observations provides well-known parameter estimates.[16] As a result of these considerations, the expressions for the "well-known" case (Table 1) are used in this work.

Limitations and ambiguities

The critical limit considers the false-positive rate (α) based on the BG distribution. It represents the signal level that is significantly greater than the BG and therefore considered "detected". Following this reasoning, L_C is an appropriate criterion to identify an event signal peak in sp-ICP-MS. The detection limit L_D provides an *a priori* performance metric for analytical methods. It requires assumptions about the distribution of the true positive signal and the tolerated false-negative proportion (β) to guarantee that a method is capable of detecting, *e.g.*, a type of NP with a certain diameter. If L_C was reported as the method detection limit, the mean value of the lowest detectable signal would be set equal to L_C. This would discard the bottom half of the true signal distribution below L_C and thus imply $\beta = 50\%$.[16] In sp-ICP-MS, this would underestimate the number concentration and overestimate the size with the magnitude of the bias depending on the number of data points per particle and the particle mass distribution.

While theory supports a detection threshold based on L_C, the commonly used formulae (Table 1) are more ambiguous. For L_D, their derivation relies on simplifications involving equal variances and distribution models for true signal and BG, $\alpha = \beta = 5\%$, and the Poisson approximation (${\sigma_B}^2 = \mu_B$). For the Gaussian model, this reduces the L_D formalism to computing "$3.29\sigma_B$".[16] The expression is numerically similar to the "$3\sigma_B$" criterion, which was defined like a critical limit with $\alpha \approx 0.135\%$, albeit with a different terminology.[14] Analogously, "$3.29\sigma_B$" could be interpreted as a critical limit with $\alpha \approx 0.05\%$. In essence, there is no absolute rule for setting significance levels, and the use of, for example, $\alpha = 5\%$ is a convention.[14,23,35] Kaiser chose $f = 3$ as a semi-empirical value which he found suitable for most cases as it includes a safety margin for deviations between model and data.[14] In other words, the fact that many authors report a value close to "$3\sigma_B$" does not require the L_D formalism to be causally correct, but rather indicates a well-suited numerical value. In sp- and sc-ICP-MS, the applicability of L_D is limited by the fact that the event signal distribution is not known *a priori,* and that the signal peaks cannot be described with the same distribution as the continuous BG. When specialized Poisson models are developed, *e.g.*, for non-integer time-of-flight analyzer data, a clear distinction between L_C and L_D is favorable.[13,32,36] For this work, however, eqn (5) and (6) are referred to as L_D out of convention and the focus lies on the numerical results.

Experimental

Sample preparation

C. reinhardtii culture preparation was carried out at the Institute of Plant Biology and Biotechnology of the University of Münster. *C. reinhardtii* wild-type strain cc124 was grown photoheterotrophically under continuous light conditions (30 μmol photons $m^{-2}\ s^{-1}$) in tris-acetate-phosphate (TAP) medium. TAP medium was prepared with 20 mmol L^{-1} tris, 2.5% 40× Beijerinck salts (v/v), 0.1% Hutner trace elements solution, and 1 mmol L^{-1} phosphate solution from K_2HPO_4 and KH_2PO_4 adjusted to a pH of 7 with acetic acid.[37] The trace elements solution was set aside for two weeks and filtered through Whatman paper before its use. All growth container glass flasks were acid-washed (32% (m/m) HCl) and autoclaved subsequently for sterilization. 24 h prior to the collection of cells, a concentration of 2×10^6 cells mL^{-1} was adjusted *via* dilution with TAP medium. The initial cell concentration was determined using a Fuchs-Rosenthal cell-counting chamber. Aliquots of the liquid culture were taken and transported on ice to the Institute of Inorganic and Analytical Chemistry where 1 mL of the suspension was removed and washed three times with ammonium acetate buffer (10 mmol L^{-1}, pH 7.5) *via*

centrifugation (4 min, 125 g, 4 °C). The pellet was resuspended in 1 mL of ammonium acetate buffer. Prior to sc-ICP-MS analysis, the washed suspension was diluted 20× (v/v) with ammonium acetate buffer to a final cell density of approximately 4×10^5 cells mL^{-1}. The cell density was determined using an automated hemocytometer without staining. Cell suspensions were handled in 2 mL microcentrifuge tubes (Eppendorf SE, Hamburg, Germany). In all cases, doubly distilled (dd.) water purified with an Aquatron A4000D water still system (Barloworld Scientific, Nemours, France) was used. For method development, *Saccharomyces cerevisiae* was obtained as a block of fresh compressed yeast from a grocery store. 100 mg of the compressed yeast were suspended in 50 mL dd. water. An aliquot of 1 mL was washed three times with dd. water *via* centrifugation (4 min, 125 g, 4 °C) and finally resuspended in 1 mL of dd. water. Prior to sc-ICP-MS analysis, the suspension was diluted to a final cell density of approximately 4×10^5 cells mL^{-1} using dd. water. For cell density measurements, 10 μL of *S. cerevisiae* suspensions were mixed with an equal volume of trypan blue stain (0.4%; Invitrogen, Carlsbad, CA, USA) and analyzed with an automated hemocytometer. Gold nanoparticles (AuNP) with a nominal size of 50 nm and 60 nm (nanoComposix, San Diego, CA, USA) were sonicated and diluted with dd. water to an approximate concentration of 4×10^5 NP mL^{-1} prior to analysis. All samples were prepared and analyzed with ICP-MS on the same day as quickly as possible.

Instrumentation

Prior to ICP-MS analysis, cell densities were monitored off-line with a Countess automated hemocytometer (Invitrogen, Carlsbad, CA, USA). A 1750 TLLX 500 μL glass syringe (Hamilton Bonaduz AG, Bonaduz, Switzerland) driven by a KDS-100 syringe pump (KD Scientific Inc., Holliston, MA, USA) was connected to an iCAP TQ ICP-MS (Thermo Fisher Scientific, Bremen, Germany) triple quadrupole ICP-MS system operated with Qtegra 2.10 (Thermo Fisher Scientific) in time-resolved analysis mode. The pump was placed in an upright position with the syringe outlet facing downward to prevent sedimentation at the moving plunger and to increase the overall distance between cells and closest surface in the direction of sedimentation. The sample flow rate was 5 μL min^{-1}. The ICP-MS system was equipped with a CytoNeb 50 low-flow concentric nebulizer, a CytoSpray total consumption on-axis spray chamber (Elemental Scientific, Omaha, NE, USA), and a nickel sampler and skimmer cone with a "high sensitivity" skimmer insert (Thermo Fisher Scientific). A one-piece torch with an injector of inner diameter of 2.0 mm (Elemental Scientific) was installed. A DPC17 external mass flow controller (Aalborg Instruments & Controls NY, USA) was used to deliver a constant Ar flow of 0.2 SLPM (standard liter per minute) to the nebulizer. The spray chamber sheath gas was supplied with approximately 1.15 SLPM of Ar from the ICP-MS instrument. Together with the interface lenses and torch position, the sheath gas flow rate was optimized daily for highest signal intensities with an oxide ratio (CeO/Ce) lower than 1.3%. Further relevant ICP-MS parameters were plasma power of 1550 W, sampling depth at 10 mm, collision/reaction (CR) cell gas He flow rate of 0.74 mL min^{-1} and O_2 flow rate of 0.25 mL min^{-1}. The CR gases were activated for all monitored elements, but oxide mass shift reactions were only used for Fe and P as $^{56}Fe \rightarrow {}^{56}Fe^{16}O$ and $^{31}P \rightarrow {}^{31}P^{16}O$. For all experiments, a dwell time of 1 ms was set and a sampling time of 60 s was used. The sample introduction system was rinsed manually between samples using dd. water first, and then 2% (v/v) nitric acid to prevent the acid from mixing with cells in the syringe, sample line or nebulizer.

Data analysis

The instrument raw data were obtained as a time-resolved count rate in units of counts per second (cps) from the manufacturer instrument software export. For the statistical analysis, and the Poisson model in particular, all data were converted to counts per dwell time (cts DT^{-1}) based on the time stamps and count rate intensities. This also allows for an easier calculation of the event summed intensity with the unit counts per event. The unit conversion together with the background signal statistics, event search and gate filter, which are described in the following, were all implemented in a Java-based (OpenJDK 15, Oracle, Austin, TX, USA) in-house software tool (single particle Tool, spTool2). A preliminary version of spTool2 has been used and tested internally and will be made publicly available. The underlying algorithms are explained in this publication. Inverse cumulative distribution functions (CDF) and empirical CDFs were calculated using MATLAB R2022a (The MathWorks, Natick, MA, USA). Pseudorandom numbers were generated with excel (Microsoft 365, Microsoft, Redmond, WA, USA), and all graphs shown were created using OriginPro 2023 (OriginLab, Northampton, MA, USA).

Results and discussion

The main focus of this work was the proposal of a modular workflow that could improve data analysis in sp- and sc-ICP-MS through separation of tasks. The statistical models use different approaches to estimate the BG distribution than for the subsequent event search. In addition, a gate filter is applied to remove falsely picked events and reallocate them to the background data. The aspect of reallocation and revoking the status as an event distinguishes this from filters that select a certain range in an event population. In the following examples, the respective elements Mg, P, Mn, Fe, Cu, Zn and Au were selected to cover different BG signal levels for the statistical analyses. The observations are not element-specific but depend on the signal intensity. Reporting the isotopic mass or m/z only serves as a reference to the examples in figures and tables within this work.

Baseline

The background signal distribution is defined based on the time-resolved raw data of each sample. Two aspects were found to improve the performance substantially. First, both a Gaussian and a Poisson model were evaluated. In Fig. 1, three examples are shown for different BG levels μ_B of 0.1 cts DT^{-1},

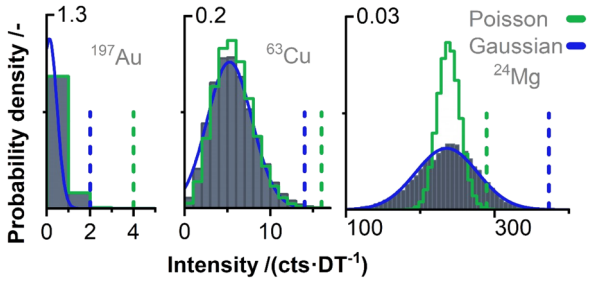

Fig. 1 Three experimental BG distributions for different mean signal levels ($\mu_{Au} = 0.118$ cts DT^{-1}, $\mu_{Cu} = 5.24$ cts DT^{-1}, $\mu_{Mg} = 235.4$ cts DT^{-1}) are plotted as signal intensity histograms (gray). A Poisson (green) and Gaussian (blue) distribution is superimposed. Dashed lines indicate Y_D ("3.29σ_B"). The Gaussian fails to describe the low intensity distribution (Au) but compensates the overdispersion at high signal levels (Mg).

5.0 cts DT^{-1}, and 235.4 cts DT^{-1}, respectively. For very low BG levels, the signal distribution has a pronounced positive skew. As a result, the Gaussian model provides a poor fit and underestimates the detection threshold. For the mid-range, both models agree with the observations to a similar degree. However, at high background levels, the experimental data are overdispersed, causing the Poisson model to underestimate the detection threshold. In previous studies, such extra-Poisson variation was explained by flicker noise contributions from the sample introduction (pump system, nebulizer, and aerosol transport) and the plasma. Solutions have been reported that include a correction term based on additional experiments.[10,36] Similarly, a doubly stochastic Poisson model or an instrument-specific compound Poisson model could be considered.[13,32,38,39]

However, for this workflow, the overdispersion at high signal intensities is addressed with a Gaussian model. The key consideration was that the model with a better fit predicted a higher threshold in all studied cases. Therefore, a data-dependent "pessimistic" decision criterion was defined, in which both models are evaluated and the worse (*i.e.*, higher) threshold is chosen. The obvious advantage is that the empirical SD of the data directly incorporates the excess variance without further experiments or modelling, while the Poisson model covers the lower signal range. There are recommendations to switch from Gaussian to Poisson when μ_B is lower than 10 cts DT^{-1} based on flicker noise analysis.[36] The "pessimistic" criterion represents an alternative that includes the *in situ* experimental conditions of a sample, which in the case of cell suspensions is difficult to simulate by a blank. In this work, the criterion was applied to a total of 668 time-resolved data sets of AuNP (50 nm), *C. reinhardtii* and *S. cerevisiae* suspensions. The Gaussian model was preferred for 124 sets, with 95% of all cases having a μ_B greater than 15.8 cts DT^{-1}. The lowest μ_B for a Gaussian model was 10.5 cts DT^{-1}. For the remaining 544 sets, the Poisson model was chosen. The highest observed μ_B for a Poisson model was 69.1 cts DT^{-1} with 95% being less than 12.1 cts DT^{-1}. In summary, the "pessimistic" criterion agrees well with the transition point of 10 cts DT^{-1} but allows for different decisions for individual samples.

Table 2 Statistical parameters (mean μ_B, standard deviation SD and detection limit Y_D) of the estimated background for two samples of *C. reinhardtii* (^{24}Mg, ^{63}Cu, DT = 1 ms) based on the initial raw data, an iterative Gaussian "$f_o \times \sigma_B$" criterion, the Poisson model in eqn (6) and the generalized extreme studentized deviate (GESD) test for outliers

	^{24}Mg [cts DT^{-1}]			^{63}Cu [cts DT^{-1}]		
	μ_B	SD	Y_D	μ_B	SD	Y_D
Raw data	322	748	2783	6.3	12	45
$f_o = 1$	239	0.0	239	5.1	0.0	6
$f_o = 2$	239	28.9	334	4.7	1.9	12
$f_o = 3$	236	40.8	370	5.1	2.4	14
GESD	236	42.1	374	5.2	2.6	14
Poisson	224	34.2	276	5.2	2.5	16

The second improvement addresses the outlier test. Matching a blank sample with a cell suspension is problematic since the analyte concentration is subject to unpredictable variations caused by the release of ions or adherent residues from the culture medium. Instead, outlier tests can remove the particle and cell signals, essentially reconstructing the underlying blank signal within each sample. This is crucial for the parameterization and comparison of the Gaussian and Poisson models. In sp- and sc-ICP-MS, the outlier identification is complicated by three main facts. First, the number of outliers is expected to be large but not known *a priori.* Second, the raw data consist of two different signal types (BG and events) whose mixed distribution function is not evident. Third, the mean and SD of the data are biased by the outliers themselves. For example, in Table 2, the SD of the raw data is an order of magnitude larger than the SD of the BG without outliers. All of this can lead to an overestimation (swamping) or underestimation (masking) of the true number of outliers, especially when tests designed to identify a single outlier are repeated.[40,41] This also concerns the "3σ" criterion and the standard normal *z*-score in general. The generalized extreme studentized deviate (GESD) test described by Rosner as a modification of Grubbs' test is reported to be more robust and used for a reference point in this work.[42,43] However, it requires normally distributed data, and the large number of outliers in sp- and sc-ICP-MS data makes it computationally intensive as it calculates the critical value and test statistic for each potential outlier. Therefore, the main focus in this work was a detailed investigation of the faster "$f_o \times \sigma_B$" outlier test to understand when swamping and masking occur in sp-ICP-MS data and how the algorithm can be stabilized. The unique characteristics of outlier statistics emphasize that the factor for the outlier test (f_o) and for the detection threshold (f_t) should be defined separately.

In Fig. 2, panel A, the "$f_o \times \sigma_B$" test is applied to time-resolved data (^{24}Mg) of a *C. reinhardtii* sample, and the remaining data points after each iteration are shown. For $f_o = 3$, both a visual inspection and a comparison with the results of the GESD listed in Table 2 support that a valid estimate of the BG is obtained. For $f_o = 1$, however, the "1σ" range with the tolerated non-outlier values around the mean is so narrow, that all data points except one numeric value ($\mu_B = 238.6$ cts DT^{-1}) are removed. The corresponding SD is zero. Consequently, this

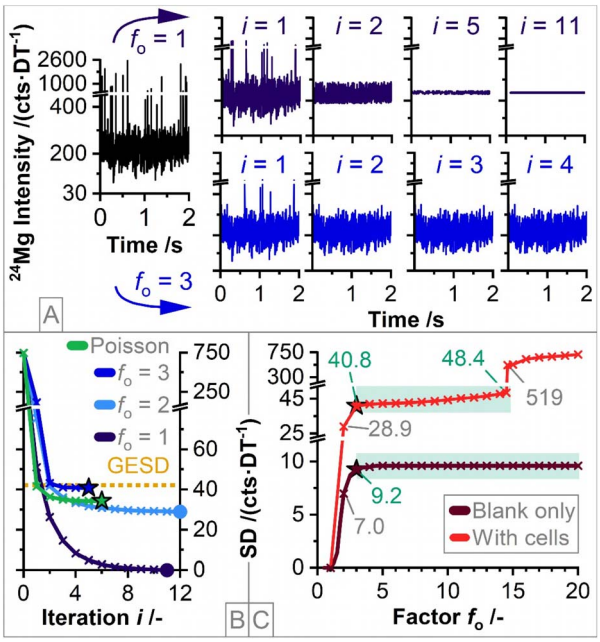

Fig. 2 Panel (A): iterative removal of outliers is performed on time-resolved raw data from a sample of *C. reinhardtii* (^{24}Mg, DT = 1 ms). At the top, $f_o = 1$ ("1σ_B", dark blue) removes all data points except for one numeric value. After $i = 11$ iterations, SD = 0 is obtained. Below, $f_o = 3$ ("3σ_B", light blue) succeeds within four iterations. Panel (B): the graph shows the SD of the same data set after each iteration for $f_o = 1$ to 3. For comparison, the result of the GSED test and the Poisson model eqn (6) are shown. Star symbols (★) mark the stopping point of the algorithm in all panels. Panel (C): the effect of different factors f_o on the SD is shown for a blank solution, and a *C. reinhardtii* sample (^{24}Mg, DT = 1 ms) using Gaussian models. The green boxes indicate the region of stable convergence that is limited by the transition points where the algorithm either underestimated the SD (swamping) or overestimates it if cell events are present and f_o is large (masking).

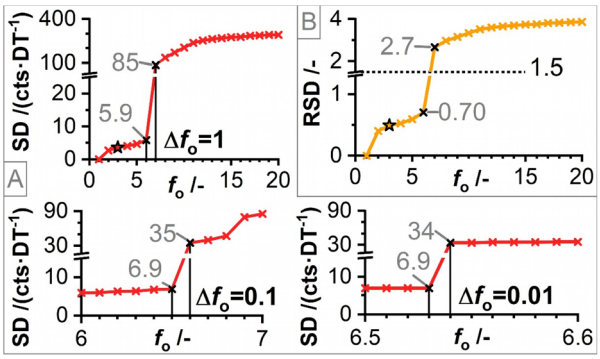

Fig. 3 Panel (A): varying the step size Δf_o between the outlier test factors f_o reveals the discrete onset of the retention of outliers in the data occurring between $f_o = 6.53$ and $f_o = 6.54$. Panel (B): at the transition, the relative standard deviation (RSD) increases above a value of RSD = 1.5. The data set is taken from a sample of *S. cerevisiae* (^{31}P → ^{31}P^{16}O, DT = 1 ms).

kind of swamping effect causes the test to estimate $\sigma_B = 0$, which is incorrect and does not provide a meaningful detection threshold with the "$f_t \times \sigma_B$" criterion. On the other hand, these observations were interpreted in such a way that the SD serves as an indicator of the goodness of the test. Thus, Fig. 2, panel B, shows the SD of the data after each iteration, including the result of the GSED test for comparison. The GESD is expected to be accurate, since the corresponding set of ^{24}Mg data was found to be normally distributed in Fig. 1. While $f_o = 1$ yields SD = 0, and $f_o = 2$ returns a non-zero SD, the swamping effect seems to be sufficiently suppressed only by incrementing again to a value of $f_o = 3$. For this value, the final SD agrees with the GESD test (Table 2). In addition, in Fig. 2, panel C, both the cell sample and blank SD only end up in the region with relatively constant SD and stable convergence when the additional increment is included. This is supported by an additional data set for ^{63}Cu data also from a *C. reinhardtii* sample, which is summarized in Table 2.

Next, the test performance with larger values of f_o in the range from 1 up to 20 was investigated. In Fig. 2, panel C, ^{24}Mg data for a sample of *C. reinhardtii* and a blank are shown. As expected, the test with $f_o = 1$ yields a SD of zero for both. For larger f_o, the blank trivially converges to the SD of the raw data, that does not contain any cell event signal (*i.e.*, outliers). When cells are present, the SD initially develops similarly, albeit at a slow, non-constant rate of increase. However, between $f_o = 14$ and $f_o = 15$, a discrete transition occurs. The SD gets substantially larger, indicating the onset of masking effects and the retention of outliers in the data. In Fig. 3, panel A and C, the increment step size Δf_o was reduced to evaluate a sample of *S. cerevisiae* based on the ^{31}P → ^{31}P^{16}O signal. Here, the masking effects start earlier, between $f_o = 6$ and $f_o = 7$. Panel C shows that even with a small increment size of $\Delta f_o = 0.01$, the transition is still discrete, perhaps because the count intensity data are based on integers.

A value of $f_o = 7$ is within the range of commonly reported values for f.[27,30] Therefore, a prevention strategy was developed. A comparison between different samples and signal intensities is simplified when the SD is normalized to the mean, *i.e.*, the relative standard deviation (RSD) calculated as RSD = SD × μ_B^{-1} is used. In Fig. 3, panel B, the RSD exhibits a sharp transition from 0.7 to 2.7 for $\Delta f_o = 1$. For this example, ^{31}P → ^{31}P^{16}O data from a sample of *S. cerevisiae* was used. Based on more observations, the comparability of the RSD between different samples and signal intensities was confirmed, and the maximum tolerated RSD for the algorithm was set to be 1.5. However, there is an exception to this rule at low BG levels. For example, Fig. S1,† panel B, shows that the RSD evolves in a range between 3 and 6 for a data set of AuNP with $\mu_B = 0.1$ cts DT^{-1}. This apparent contradiction with the previous statement arises from the property of the Poisson distribution, that is $\sigma = \sqrt{\mu}$. For values of $0 < \mu_B < 1$, it follows that $\sqrt{\mu_B} > \mu_B$ which explains why the RSD is greater than one for all values of f_o. Although a Gaussian model is not ideal for $\mu_B < 10$, it is not uncommon in the literature and also required for the comparison of Gaussian and Poisson models in this work.[28,29] In fact, for $f_o = 5$, a variance-to-mean ratio (VMR, $\sigma_B^2 \times \mu_B^{-1}$) of 0.92 indicates that the Gaussian outlier test generates a reasonable BG data set that is consistent with the expected VMR (VMR = 1) for Poisson distributed data.

It should be noted that the point of transition depends on the event signal frequency and intensity relative to the background, as well as the number of points per event. Thus, for different experimental conditions, the corresponding value of f_o is difficult to predict, giving another indication as to why a wide range of values has been reported for f elsewhere.[30,44] As summarized in Table S1,† samples with a low BG signal like AuNP usually require a factor of $f_o \approx 5$. The onset of masking effects for the sample of *S. cerevisiae* ($^{31}P \rightarrow {}^{31}P^{16}O$) occurred at $f_o = 6.6$. In summary, the lower limit of f_o for one kind of sample may be close to the upper limit of f_o for another sample. This makes it difficult to propose a universal choice for f_o. Considering that f_o does not express a significance level in the true sense when used in a repeated test, the algorithm described in the following section was developed to simply find a "working value" of f_o.

The algorithm is illustrated as a flowchart in Fig. 4. Here, a brief summary is given. The analysis starts with $f_o = 1$ and increments by $\Delta f_o = 1$ until a non-zero SD is obtained. The algorithm interprets SD < 0.05 as "zero". This was found to be sufficient and more stable with respect to floating point arithmetic paired with imprecision due to a lack of significant figures in the manufacturer raw data csv-files. In the shown example, $f_o = 2$ was enough (Fig. 2, Table 2). The first successful value for f_o is incremented by $\Delta f_o = 1$, in this case to $f_o = 3$, as already discussed. Next, the algorithm calculates the RSD. For $\mu_B < 1$, RSDs greater than 1.5 are tolerated. If the mean is greater than one ($\mu_B > 1$), the algorithm requires RSD < 1.5. Otherwise, it will decrement the value of f_o by one and start incrementing using steps of $\Delta f_o = 0.1$. That means, if the algorithm yields SD ≥ 0 for $f_o = 4$ and RSD > 1.5 for $f_o = 5$, it will start incrementing at 4.1σ, 4.2σ, … < 5σ. Since the transition occurs in a point-like manner even at step sizes of $\Delta f_o = 0.01$, choosing $\Delta f_o = 0.1$ was found to be sufficient to stop before the point is reached. For the graph shown in Fig. 2, $f_o = 14.5$ corresponds to SD = 48.4 cts DT^{-1} whereas $f_o = 14.6$ already leads to SD = 519 cts DT^{-1}. Here, $f_o = 14.5$ would be chosen.

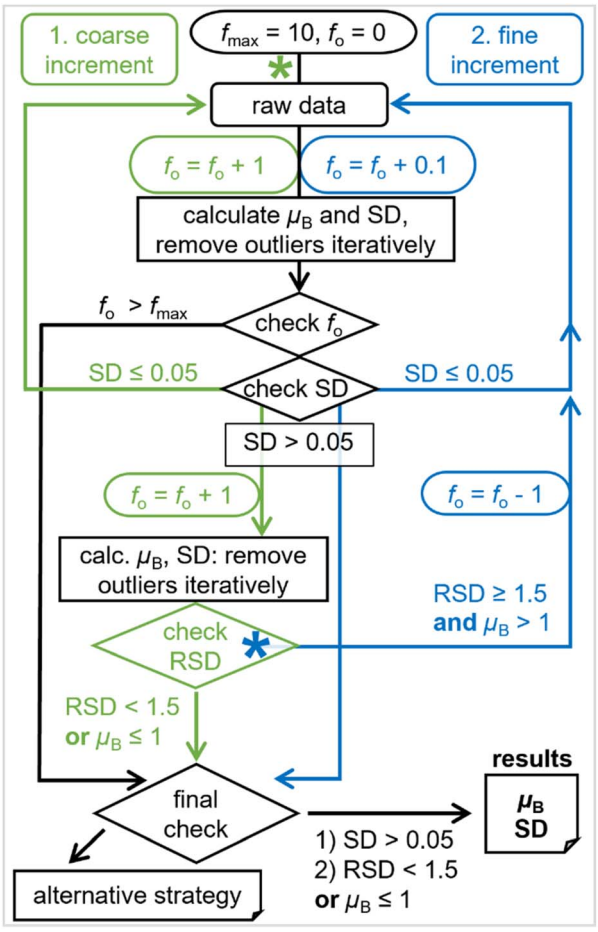

Fig. 4 The flowchart describes the algorithm for the improved, incremented removal of outliers. Asterisk symbols (*) denote the starting points for the coarse and fine step size, respectively. Note that the values for the check conditions $\mu_B > 0.05$ as well as RSD < 1.5 are empirical and may require adjustment. If the algorithm does not succeed, an alternative strategy should be chosen. This could be a Poisson approximation based on either the raw data median, the last value of μ_B or a default value such as $\mu_B = 0.1$.

As a final recommendation, it should be defined how the algorithm proceeds when the raw data have such a low BG level that virtually no result other than SD ≈ 0 can be obtained. In this case, using a normal approximation to Poisson with $\sigma^2 = \mu$ is a suitable choice. As a value for μ, both the mean of the last iteration, but also the median of the non-zero raw data or a default value $\mu = 0.1$ were successfully tested. If the "pessimistic" criterion is applied subsequently, the Poisson model usually provides the higher detection threshold and replaces the intermediate result of the unsuitable Gaussian model. All in all, this set of instructions was found to perform autonomously under various test scenarios and for all data sets involved in this work.

Interestingly, as shown in Fig. 2, panel B, an iteration based on the Poisson model in eqn (6) cannot converge to zero because of the constant summand of 2.71, which is a useful feature of that model in the context of outlier removal. If eqn (3) is used, the swamping effect occurs due to the low value of f_o. Thus, using the corresponding Y_D expression of eqn (4) is recommended for the outlier test.

Event search

Implementing the modular principle, the outlier test statistics were treated separately from the detection threshold and *vice versa*. In this work, "3.29 σ_B" was used as the detection threshold corresponding to the Y_D-type formulae in Table 1, regardless the value of the factor f_o in the outlier test. The Y_C-type formulae were considered but the difference was negligible when the gate filter, which will be explained next, was used. In sp-ICP-MS, event signal durations of 0.5 ms and longer have been reported.[33,34] At a DT of 1 ms, the split-event probability becomes >50%. Thus, a split-event correction was applied, combining all consecutive data points above the detection threshold into a single event. The calculated detection thresholds were rounded to the higher integer value.[36,45] Consequently, the search algorithm in this work considered all detector readings greater than or equal to the threshold as part of an event.

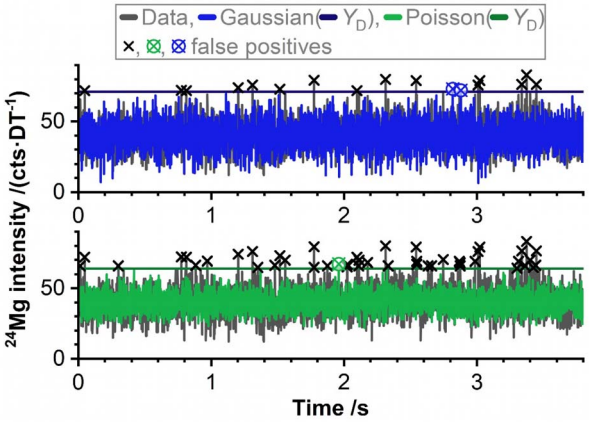

Fig. 5 A section of time resolved data (^{24}Mg, DT = 1 ms) from a blank sample (gray) is overlayed with pseudo-random numbers from a Gaussian (blue) and Poisson (green) distribution, respectively. Their parameters (μ_B, SD) are estimated from the raw data. Cross symbols (×) indicate points above Y_D, revealing an excess of false-positive signals.

Gate filter

Despite the compensation of extra-Poisson variance with a Gaussian model as presented in the first data processing step, observations of excess false-positive signals persisted. To visualize this, in Fig. 5, time-resolved raw data of an ammonium acetate buffer blank sample are overlayed with simulated pseudorandom numbers from a Gaussian and Poisson distribution based on the sample mean intensity and SD. In comparison, the raw data have an unexpectedly high number of signals above Y_D from eqn (4)–(6). Extra-Poisson variance explains why the frequency of these extreme events is larger for the Poisson ($N = 45$) than for the Gaussian ($N = 15$) model. However, it cannot explain why the respective cumulative distribution functions (CDF) deviate when the upper quantiles are examined in detail as shown in Fig. 6. Instead, the CDFs demonstrate an excess of extreme BG events outside the scope of the model. Plausible causes include changes in the flicker-noise components such as mass-flow fluctuations caused by the pump, variations in the droplet size distribution or turbulent aerosol transport in the spray chamber.[46] In sc-ICP-MS, cell-compatible buffers usually carry high salt concentrations, inducing deposits on the nebulizer and in the spray chamber with adverse effects on the aerosol stability. In addition, both cell deterioration and fixation are known to generate aggregates with the potential to temporarily change the fluid flow through the sample line and the nebulizer. On the other hand, a sample may also contain particulate contaminations. Some metal species are not stable in cell culture media or to matrix changes during the sample preparation and may form NP through precipitation. In addition, cells can release particulate species such as vesicles, aggregated macromolecules, micelles, or fragments. If the sample cleanup relies on a centrifuge, any particulate matter of similar sedimentation coefficients cannot be separated. Such a mixed BG with particulate and dissolved components is a challenge for existing detection threshold models. The particulate fraction is not part of the Gaussian or

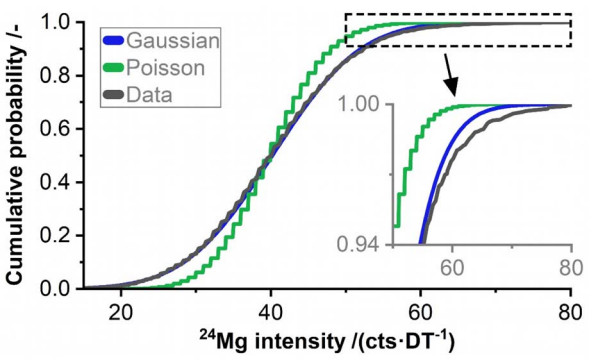

Fig. 6 The empirical cumulative distribution function (CDF) of a blank sample (gray) is superimposed with a Gaussian (blue) and Poisson (green) CDF. Their parameters (μ_B, SD) are estimated from the raw data (^{24}Mg, DT = 1 ms). While the Gaussian matches the overall trend well, the inset shows a deviation at the upper end, suggesting an excess of false-positives in the data.

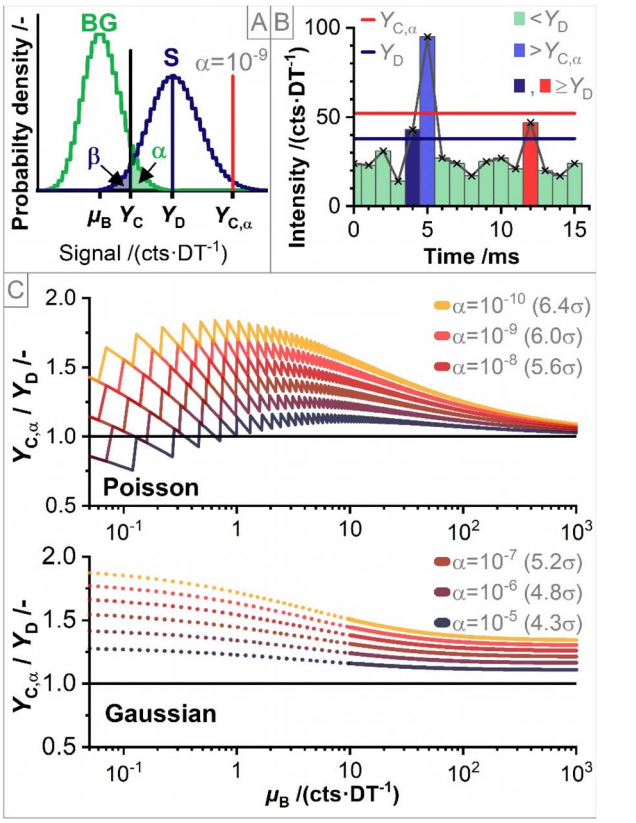

Fig. 7 Panel (A): Poisson probability densities for background (BG) and signal (S) illustrate the definitions of Y_C and Y_D for $\alpha = \beta = 5\%$. An additional critical value $Y_{C,\alpha}$ with higher significance levels is shown ($\alpha = 10^{-9}$, red line). Panel (B): exemplary cell signal peak above a pseudorandom number BG with DT = 1 ms. Y_D (dark blue) and $Y_{C,\alpha}$ ($\alpha = 10^{-9}$, red line) are shown. If Y_D was increased to filter the false-positive signal (red), the dark blue peak area would also be lost. Panel (C): the ratio r_g between the gross critical value $Y_{C,\alpha}$ and Y_D based on eqn (5) and (6) is plotted against the mean BG signal μ_B. For $\alpha = 10^{-5}$ (4.3σ) to $\alpha = 10^{-10}$ (6.4σ), the ratio predominantly takes a value between 1 and 2 for both Gaussian and Poisson models.

Poisson model and could cause unexpected small signals above the predicted detection threshold. False-positives can result from the simultaneous registration of multiple particles (coincidence) even if the individual particulate contaminants are smaller than the detection threshold. The same kind of argument can be made for sp-ICP-MS if a sample contains a heterogenous mixture of different NP sizes and number concentrations.

It would be possible to reduce the false-positive rate by choosing a higher detection threshold. However, this approach is not suitable for DTs of 1 ms and below, where multiple data points per signal event peak are obtained. The split-correction event search would incorrectly reject data points with lower intensities at the edges of the peak, as shown in panel B of Fig. 7. This would underestimate the intensity sum of the peak and the corresponding element mass of the event. Instead, in this work, a secondary threshold is introduced as a gate filter. It removes all initially identified events whose peak height is below the gate threshold. If the peak height exceeds the gate threshold, the start and end points of the peak as well as the signal sum are not affected. As the main advantage, this secondary filter stage can be fine-tuned to address the issue of excessive false-positive events without changing the peak area of the valid events. The peak height was selected as the criterion for three reasons. First, the BG statistics predict the signal per DT and not, *e.g.*, the intensity per cell event. Analogously, the number of data points per event, which strongly depends on the DT, was discarded. Second, a peak height filter rewards better signal-to-background ratios, providing a positive indicator for method development. Third, a key result in sp- and sc-ICP-MS analyses is the summed intensity of an event and the shape of the distribution in a histogram. When the gate filter evaluates a different parameter (height instead of area), it causes less distortion of the event intensity distribution than, for example, cutting out a narrower range in the histogram.

Since the cause and underlying distribution of the false-positive events are not known, their frequency is difficult to predict. Extrapolating the frequency from a blank sample would require both a suitable cell suspension blank, and a long sampling time to observe enough rare false-positive events. Therefore, the preference was to establish an appropriate gate threshold based on the BG distribution, which is estimated for each sample using outlier tests. A first approach was developed by applying different empirical gate filter thresholds Y_{gate} to an event signal distribution that is sufficiently separated from the BG to visualize the change, as shown in Fig. 8 and Fig. S2.† The gate threshold Y_{gate} was adjusted until most false-positive signals in the tail of the BG distribution and below the core distribution were removed. This procedure was tested with a variety of samples of different cells, buffers, NPs, and BG signal levels. It was observed that a suitable gate filter value can be calculated by multiplying the gross detection limit Y_D with a constant factor f_g, as shown in eqn (7). In this work, a gate factor of $f_g = 2$ was found to perform well.

Further considerations support this approach. The false-positive events are located in the top quantiles of the raw data. Consequently, the distributions have longer tails than the Gaussian or Poisson distribution predicts (Fig. 6). The gate threshold Y_{gate} can be placed in these top quantiles by calculating its value as a critical limit $Y_{C,\alpha}$ with an exceptionally low α-value as shown in eqn (8). To evaluate the gate factor approach with $f_g = 2$, the ratio of $Y_{C,\alpha}$ and the gross detection limit Y_D was analyzed. The ratio is referred to as r_g and is expected to be of similar magnitude as f_g. According to eqn (9), r_g was calculated

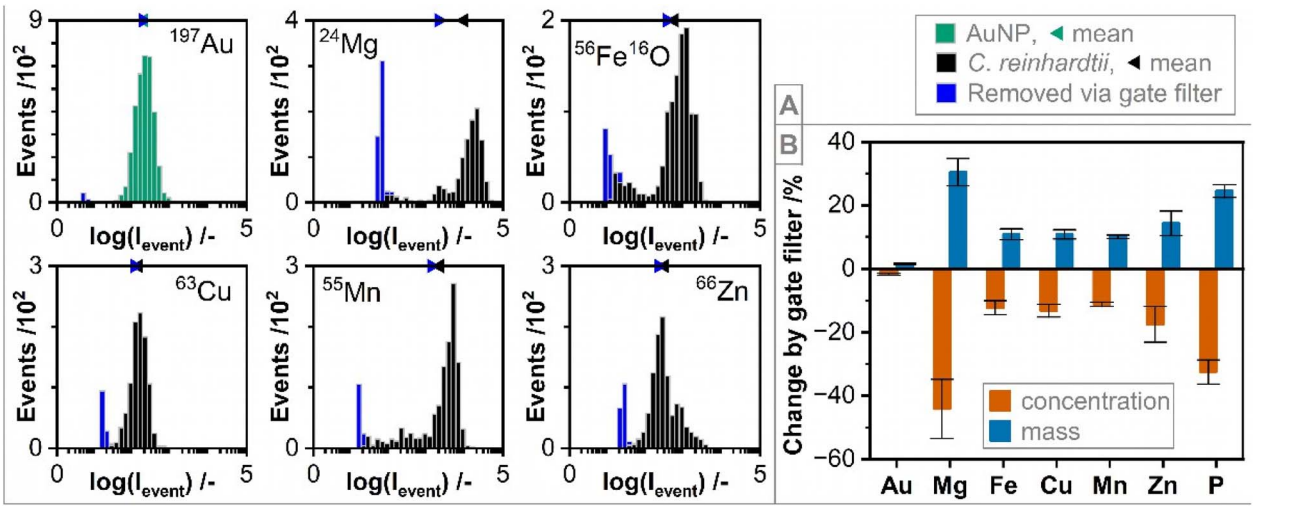

Fig. 8 Panel (A): histograms show the event signal distribution after $\log_{10}$-transformation (DT = 1 ms). The gate filter ($\alpha = 10^{-10}$) identifies more false-positive events (blue) in *C. reinhardtii* sc-ICP-MS data compared to AuNP. The detection thresholds (Y_D expressions) are based on the optimized outlier test and the "pessimistic" criterion. Triangular symbols (►, ◄) indicate the mean of the entire data set (blue) and the remaining data after applying the gate filter (black). Panel (B): the cell and NP concentration (cells mL^{-1}, NP mL^{-1}) is calculated from the number of events, and the element mass per event from the signal, respectively. The difference of each parameter before and after applying the gate filter is normalized to the value of the filtered data ($N = 3$). For *C. reinhardtii*, the unfiltered data would underestimate the mass per cell up to 30% (17% on average) and overestimate the cell concentration up to 44% (22% on average). For AuNP, the effect was only 2% for both parameters.

as a function of μ_B for α going from 10^{-5} to 10^{-10} which corresponds to 4.3σ and 6.4σ, respectively, in terms of a normal distribution. Since the false-positive events are not part of the model, these α-values should not be treated as strict significance levels, but as a computational tool. First, the Gaussian model will be discussed. It requires σ_B to calculate both Y_D *via* eqn (5) and $Y_{C,\alpha}$ which can be written as eqn (10). In order to account for extra-Poisson variance and flicker noise, σ_B was calculated as a function of μ_B using a quadratic polynomial as described elsewhere.[10,36] To cover different values of μ_B, BG data for varying m/z but the same experimental setup were fitted with a quadratic polynomial whose constant offset was set to zero and linear coefficient was set to one ($\sigma_B{}^2 = \mu_B + 0.031 \times \mu_B{}^2$, $R^2 = 0.9987$). Since Gaussian models work best for $\mu_B > 10$ cts DT^{-1}, the evaluation of r_g was limited to this range.[36] Depending on the values of α and μ_B, ratios between $r_g = 1.1$ and $r_g = 1.5$ were obtained (Fig. 7).

For the Poisson model, $Y_{C,\alpha}$ was calculated from μ_B using the inverse CDF. For Y_D, eqn (6) was used. Ratios of up to $r_g = 1.8$ were obtained (Fig. 7). For large μ_B, the limit of r_g tends towards one. However, it is expected that a Gaussian model will be selected based on the "pessimistic" criterion or the "$\mu_B > 10$ cts DT^{-1}" criterion before this point is reached. For small μ_B, the ratio r_g takes a pronounced skewed step pattern because the Poisson $Y_{C,\alpha}$ only takes integer values. The smallest meaningful value of $Y_{C,\alpha}$ ($Y_{C,\alpha} = 1$) sets the lower boundary for the ratio r_g at $Y_D{}^{-1} \approx 2.71^{-1} \approx 0.369$. The summand 2.71 in the L_D approximation eqn (6) comes from the heteroskedasticity of the Poisson distribution, and a constant of similar value results in exact Poisson calculation as well. Thus, r_g does not diverge when μ_B approaches zero. In addition, it will take a value roughly between $r_g = 1$ and $r_g = 2$ for the BG levels μ_B in most applications. In the 668 samples analyzed in this work, for example, μ_B ranged from 2×10^{-3} cts DT^{-1} to 369.1 cts DT^{-1} and the corresponding r_g were between 1.3 and 1.8 for $\alpha = 10^{-10}$. The respective distribution model was chosen *via* the "pessimistic" criterion. In summary, these considerations suggest that the assumption $Y_{gate} \approx f_g \times Y_D$ with $f_g = 2$ is a plausible choice, as f_g should at least match the highest value of r_g. Moreover, the approach is easy to implement and customize *via* f_g. The more accurate alternative would be to calculate the inverse CDF for each sample at a given α-value. In this case, the empirical SD of the BG should replace the estimation of σ_B *via* the quadratic polynomial for the Gaussian model.

$$\text{Approach 1: } Y_{gate} \approx f_g \times Y_D \tag{7}$$

$$\text{Approach 2: } Y_{gate} = Y_{C,\alpha} = \mu_B + L_{C,\alpha} \tag{8}$$

$$r_g = Y_{C,\alpha} \times Y_D{}^{-1} \tag{9}$$

$$Y_{C,\alpha} = \mu_B + z_{1-\alpha} \times \sigma_B \tag{10}$$

Application

Finally, the effects of the gate filter on particle and cell number concentration (PNC) and mass per particle and cell, respectively, are discussed. PNC, and mass per particle and cell are important target quantities in sp- and sc-ICP-MS. The PNC is calculated from the number of events, and the mass from the event signal. A change in these parameters propagates directly and linearly to the PNC and mass. Therefore, a statement about the relative effects of the gate filter can be made without a dedicated calibration. For this purpose, the difference of the data sets before and after applying the gate filter was normalized to the filtered value. In Fig. 8, the normalized change is shown with the corresponding histograms for a gate filter $\alpha = 10^{-10}$. The false-positive events have a lower mean intensity than the true events. In the unfiltered data, the false-positives therefore lead to an underestimation of the mass and overestimate the PNC. The gate filter counteracts this bias. For AuNP, the effect was small and on the order of 2%. For *C. reinhardtii* samples, the gate filter increased the estimated mass per cell by an average of 17% (up to 30% for ^{24}Mg) and reduced the assumed PNC by an average of 22% (up to 44% for ^{24}Mg). The degree of bias was greater for the cell samples than for AuNP, but the general BG signal was also higher in these samples. Therefore, it is not possible to conclude whether the cause is specific to the cell suspensions. Nonetheless, these examples emphasize the importance and benefits of data-dependent processing strategies. In addition, the cellular Mn content was quantified using an external calibration for Mn and the transport efficiency (TE). The TE was determined with AuNP based on the particle frequency (TE = 0.40). For six technical replicates, a mean Mn content of $(2.6 \pm 0.2) \times 10^7$ atoms per cell was obtained which is in accordance with the literature reporting a range of 2×10^7 to 4×10^7 atoms per cell for Mn in *C. reinhardtii*.[47]

Conclusions

In this work, the advantages of a modular workflow were shown. The reliability of the widely-used iterative "$f_o \times \sigma_B$" outlier test for Gaussian background (BG) models was improved by monitoring the standard deviation (SD) of the data set while performing the test. In short, if the obtained SD is almost equal to zero, the chosen test factor f_o is too small and the number of outliers is overestimated (swamping effect). If, however, the relative SD (RSD) is greater than 1.5, the test fails to reject all outliers and f_o needs to be reduced (masking effect). Since it was found that the transition into masking effects depends on the decimal place of the outlier factor, non-integer step sizes can be chosen as increments to solve this problem. These observations led to the conclusion that the "$f_o \times \sigma_B$" outlier test has two critical tipping points with respect to f_o. When f_o is close to these points, a small variation of its value will have a strong and unpredictable impact on the estimated BG parameters. Thus, it is crucial that the factor for the outlier test f_o should be chosen independently of the factor for the "$f_t \times \sigma_B$" criterion that sets the detection threshold, *e.g.*, as "$3\sigma_B$" or "$3.29\sigma_B$". For the Poisson model, using the "well-known"-case detection limit L_D defined by Currie[16] as an outlier test threshold was found to be ideal since the constant summand of 2.71 in the expression intrinsically supports the convergence. It was shown that

these measures enable a better comparability of the Gaussian and Poisson model thresholds. Following this idea, a selection criterion was established by the fact that the model with a better fit to the data yielded the higher detection threshold. Importantly, the Gaussian approach naturally accounts for extra-Poisson variance from flicker noise. However, a Poisson model is excellent for low count rates where a Gaussian model suffers from the skewness and integer nature of the ICP-MS count data, neither of which can be represented by the model. Furthermore, a remaining excess of false-positive events, observed even in blank samples, was discussed, and addressed with a gate filter. Since increasing the detection threshold comes with the risk of falsely excluding the edges of a split-event peak, a secondary filter based on the peak height, that does not alter the peak area, was chosen. Empirically, it was established that doubling the gross detection limit Y_D yields a sufficient gate filter threshold. Further investigation concluded that critical values with exceptionally low false-positive tolerances provide a more flexible calculation of the proposed gate filter. Further evidence that doubling Y_D indeed leads to a filter threshold with the right order of magnitude was obtained as well. An important conclusion was that single-particle and single-cell ICP-MS data should be treated with special care when approaching the detection limit. In conclusion, a modular combination of these steps was able to automatically process data of *C. reinhardtii* samples in a data-dependent manner. On average, the event count used to estimate the cell concentration was overestimated by 22%, while the event signal, which is proportional to the element mass per cell, was underestimated by 17% without a correction such as the presented gate filter. Both findings indicate that the presented approach improves upon existing data evaluation processes.

Author contributions

Matthias Elinkmann: conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, writing – original draft, Sarah Reuter: conceptualization, investigation, methodology, writing – review & editing, Michael Holtkamp: formal analysis, methodology, writing – review & editing, Steffen Heuckeroth: software, validation, writing – review & editing, Alexander Köhrer: investigation, validation, writing – review & editing, Katharina Kronenberg: investigation, methodology, writing – review & editing, Michael Sperling: conceptualization, methodology, resources, writing – review & editing, Oliver Rubner: formal analysis, validation, writing – review & editing, C. Derrick Quarles Jr: conceptualization, resources, writing – review & editing, Michael Hippler: conceptualization, funding acquisition, resources, supervision, writing – review & editing, Uwe Karst: conceptualization, funding acquisition, resources, supervision, writing – review & editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Parts of this work have been funded by the Deutsche Forschungsgemeinschaft (DFG) – CRC 1450 – 431460824.

Notes and references

1 D. Pompéu de Moraes, S. González-Morales, J. Toledano-Serrabona, M. Á. Sánchez-Garcés, J. Bettmer, M. Montes-Bayón and M. Corte-Rodríguez, *J. Trace Elem. Med. Biol.*, 2023, **77**, 127143.
2 S. B. Seiffert, M. Elinkmann, E. Niehaves, A. Vennemann, D. Mozhayeva, S. Kröger, M. Wiemann and U. Karst, *Anal. Chem.*, 2023, **95**, 6383–6390.
3 S. E. Szakas, R. Lancaster, R. Kaegi and A. Gundlach-Graham, *Environ. Sci.: Nano*, 2022, **9**, 1627–1638.
4 E. Bolea-Fernandez, A. Rua-Ibarz, M. Velimirovic, K. Tirez and F. Vanhaecke, *J. Anal. At. Spectrom.*, 2020, **35**, 455–460.
5 S. Meyer, R. G. De Vega, X. Xu, Z. Du, P. A. Doble and D. Clases, *Anal. Chem.*, 2020, **92**, 15007–15016.
6 R. Álvarez-Fernández García, M. Corte-Rodríguez, M. Macke, K. L. Leblanc, Z. Mester, M. Montes-Bayón and J. Bettmer, *Analyst*, 2020, **145**, 1457–1465.
7 A. E. Olbrich, B. A. A. Stepec, N. Wurzler, E. C. Terol, A. Koerdt and B. Meermann, *Metallomics*, 2022, **14**, mfac083.
8 E. Mavrakis, L. Mavroudakis, N. Lydakis-Simantiris and S. A. Pergantis, *Anal. Chem.*, 2019, **91**, 9590–9598.
9 S. Miyashita, A. S. Groombridge, S. I. Fujii, A. Minoda, A. Takatsu, A. Hioki, K. Chiba and K. Inagaki, *J. Anal. At. Spectrom.*, 2014, **29**, 1598–1606.
10 J. Tuoriniemi, G. Cornelis and M. Hassellöv, *J. Anal. At. Spectrom.*, 2014, **29**, 743–752.
11 D. Mozhayeva and C. Engelhard, *J. Anal. At. Spectrom.*, 2019, **34**, 1571–1580.
12 T. E. Lockwood, R. Gonzalez De Vega and D. Clases, *J. Anal. At. Spectrom.*, 2021, **36**, 2536–2544.
13 A. Gundlach-Graham and R. Lancaster, *Anal. Chem.*, 2023, **95**, 5618–5626.
14 H. Kaiser, *Fresenius' Z. Anal. Chem.*, 1965, **209**, 1–18.
15 H. Kaiser, *Fresenius' Z. Anal. Chem.*, 1966, **216**, 80–94.
16 L. A. Currie, *Anal. Chem.*, 1968, **40**, 586–593.
17 L. A. Currie, *J. Radioanal. Nucl. Chem.*, 2008, **276**, 285–297.
18 R. Peters, Z. Herrera-Rivera, A. Undas, M. Van Der Lee, H. Marvin, H. Bouwmeester and S. Weigel, *J. Anal. At. Spectrom.*, 2015, **30**, 1274–1285.
19 F. Laborda, J. Jiménez-Lamana, E. Bolea and J. R. Castillo, *J. Anal. At. Spectrom.*, 2011, **26**, 1362–1371.
20 Z. Liu, A. Xue, H. Chen and S. Li, *Appl. Microbiol. Biotechnol.*, 2019, **103**, 1475–1483.
21 C. Degueldre, P. Y. Favarger and C. Bitea, *Anal. Chim. Acta*, 2004, **518**, 137–142.
22 F. Laborda, J. Jiménez-Lamana, E. Bolea and J. R. Castillo, *J. Anal. At. Spectrom.*, 2013, **28**, 1220–1232.
23 M. Tanner, *J. Anal. At. Spectrom.*, 2010, **25**, 405–407.
24 J. Noireaux, R. Grall, M. Hullo, S. Chevillard, C. Oster, E. Brun, C. Sicard-Roselli, K. Loeschner and P. Fisicaro, *Separations*, 2019, **6**, 1–13.

25 C. Degueldre, P. Y. Favarger and S. Wold, *Anal. Chim. Acta*, 2006, **555**, 263–268.
26 S. Hu, R. Liu, S. Zhang, Z. Huang, Z. Xing and X. Zhang, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 1096–1103.
27 J. Navratilova, A. Praetorius, A. Gondikas, W. Fabienke, F. von der Kammer and T. Hofmann, *Int. J. Environ. Res. Public Health*, 2015, **12**, 15756–15768.
28 S. Lee, X. Bi, R. B. Reed, J. F. Ranville, P. Herckes and P. Westerhoff, *Environ. Sci. Technol.*, 2014, **48**, 10291–10300.
29 S. Miyashita, H. Mitsuhashi, S. ichiro Fujii, A. Takatsu, K. Inagaki and T. Fujimoto, *Anal. Bioanal. Chem.*, 2017, **409**, 1531–1545.
30 J. Tuoriniemi, G. Cornelis and M. Hassellöv, *Anal. Chem.*, 2012, **84**, 3965–3972.
31 J. Tuoriniemi, G. Cornelis and M. Hassellöv, *J. Anal. At. Spectrom.*, 2015, **30**, 1723–1729.
32 L. Hendriks, A. Gundlach-Graham and D. Günther, *J. Anal. At. Spectrom.*, 2019, **34**, 1900–1909.
33 J. Fuchs, M. Aghaei, T. D. Schachel, M. Sperling, A. Bogaerts and U. Karst, *Anal. Chem.*, 2018, **90**, 10271–10278.
34 E. Bolea-Fernandez, D. Leite, A. Rua-Ibarz, T. Liu, G. Woods, M. Aramendia, M. Resano and F. Vanhaecke, *Anal. Chim. Acta*, 2019, **1077**, 95–106.
35 L. A. Currie, *Pure Appl. Chem.*, 1995, **67**, 1699–1723.
36 F. Laborda, A. C. Gimenez-Ingalaturre, E. Bolea and J. R. Castillo, *Spectrochim. Acta, Part B*, 2019, **159**, 105654.
37 S. H. Hutner, *The Chlamydomonas Sourcebook: A Comprehensive Guide to Biology and Laboratory Use*, ed. Elizabeth H. Harris, Elsevier Science & Technology, San Diego, 1990, vol. 65.
38 A. Ulianov, O. Müntener and U. Schaltegger, *J. Anal. At. Spectrom.*, 2015, **30**, 1297–1321.
39 A. Gundlach-Graham, L. Hendriks, K. Mehrabi and D. Günther, *Anal. Chem.*, 2018, **90**, 11847–11855.
40 G. L. Tietjen and R. H. Moore, *Technometrics*, 1972, **14**, 583–597.
41 S. M. Bendre, *Commun. Stat.*, 1989, **18**, 697–710.
42 F. E. Grubbs, *Technometrics*, 1969, **11**, 1–21.
43 B. Rosner, *Technometrics*, 1983, **25**, 165–172.
44 H. E. Pace, N. J. Rogers, C. Jarolimek, V. A. Coleman, C. P. Higgins and J. F. Ranville, *Anal. Chem.*, 2011, **83**, 9361–9369.
45 C. Trujillo, J. Pérez-Arantegui, R. Lobinski and F. Laborda, *Nanomaterials*, 2023, **13**, 1–16.
46 D. Torregrosa, G. Grindlay, L. Gras and J. Mora, *J. Anal. At. Spectrom.*, 2023, **38**, 1874–1884.
47 S. Chen, S. Schmollinger and S. S. Merchant, *Metallomics*, 2023, **15**, mfad025.

