Supplementary Information

Number of spikes in single particle ICP-MS time scans: from the very dilute to the highly concentrated range

Pierre-Emmanuel Peyneau and Martin Guillon
GERS-LEE, Univ Gustave Eiffel, IFSTTAR, F-44344 Bouguenais, France

Contents

S1 Monte Carlo simulations **2**

S2 Effect of the probability distributions on the average number of spikes **2**

S3 Normality of the spike count **3**

S4 Operating parameters **3**

S5 Comparison between spike detection algorithms **4**

S6 Single particle ICP-MS analysis of dispersions composed of a mixture of 60 and 150 nm-sized nanoparticles **5**

List of Tables

S1 Dirac, Gamma, inverse Gaussian and bimodal distributions average and variance . . . 2
S2 Shapiro-Wilk test statistic W values and corresponding p-values for the three simulated nanoparticle flux rates. 3
S3 Comparison between the mean spike counts detected with our own custom-written program and with MassHunter. 5

List of Figures

S1 Average simulated vs. theoretical numbers of spikes in a time scan for several distributions of the particle event duration 3
S2 Empirical distribution of the spike count N_s for $\lambda = 10$, 100 and $1000\,\mathrm{s}^{-1}$ 4
S3 Mean spike count as a function of the nanoparticle flux rate for a mixture of differently sized nanoparticles 6

S1 Monte Carlo simulations

We generated synthetic and simplified time scans with the following procedure:

- First, we drew randomly the points of discontinuity of a homogeneous Poisson process of intensity λ. These points model the starts of the particle events left by the ions clouds stemming from the nanoparticles in the time scan. To perform this step, a series of random interarrival times, distributed according to an exponential law of parameter λ, was generated.
- Then, we drew randomly the duration of each particle event according to one of the three distributions mentioned in Sec S2. All the results presented in the article have been obtained with a constant-valued duration. More complex probability distributions have only been used to derive the results presented in Sec. S2.
- If the intersection between a particle event and a given time bin of the synthetic time scan was not void, a nonzero value (e.g., 1) was affected to the corresponding reading. All the readings of the time scan that did not belong to a fraction of any particle event were set equal to 0.

The number of nanoparticle having entered the instrument is equal to the number of discontinuities of the homogeneous Poisson process generated during the first step of the algorithm up to a duration τ_{obs}. The spike count can be calculated thanks to the procedure presented in Secs. 3.3 and S5, with a zero noise level for the Monte Carlo time scans.
The previous three-step procedure was repeated numerous times to get the moments (average, variance) or even in some cases (cf. Sec. S3) the empirical distributions of the spike count. All the Monte Carlo results reported in this work are based on at least 5,000 trials.

S2 Effect of the probability distributions on the average number of spikes

To check that the average number of spikes depends only on the average of the duration τ of particle events and not on the shape of the probability distribution function of this random variable, we ran Monte-Carlo simulations for four possible distributions:

- Constant-valued (i.e. Dirac-distributed) $\tau = \tau_p$;
- Gamma distribution: $\text{PDF}(\tau) = \frac{1}{\Gamma(k)\theta^k}\tau^{k-1}e^{-\tau/\theta}$, where Γ denotes the gamma function – $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}\,\mathrm{d}t$ for $x > 0$;
- Inverse Gaussian (IG) distribution: $\text{PDF}(\tau) = \sqrt{\frac{\lambda}{2\pi\tau^3}}\exp\Big(-\frac{\lambda(\tau-\tau_m)^2}{2\tau_m^2\tau}\Big)$;
- A bimodal distribution such that $P(\tau = \tau_{p,1}) = 0.5$ and $P(\tau = \tau_{p,2}) = 0.5$.

Gamma and inverse Gaussian distributions are commonly employed to model nonnegative real random variables.
The relationships existing between the parameters of the distributions and their first two moments (average and variance) are given in Tab. S1.

Table S1: Dirac, Gamma, inverse Gaussian and bimodal distributions average and variance

Particle event duration PDF	Average	Variance
δ_{τ_p}	τ_{p}	0
Gamma(k, θ)	$k\theta$	$k\theta^2$
IG(τ_m, λ)	τ_m	τ_m^3/λ
$0.5(\delta_{\tau_{p,1}} + \delta_{\tau_{p,2}})$	$(\tau_{p,1} + \tau_{p,2})/2$	$(\tau_{p,1} - \tau_{p,2})^2/4$

Fig. S1 confirms that the average number of spikes only depends on the average value of τ regardless of the underlying probability distribution of τ, including when it is multimodal.

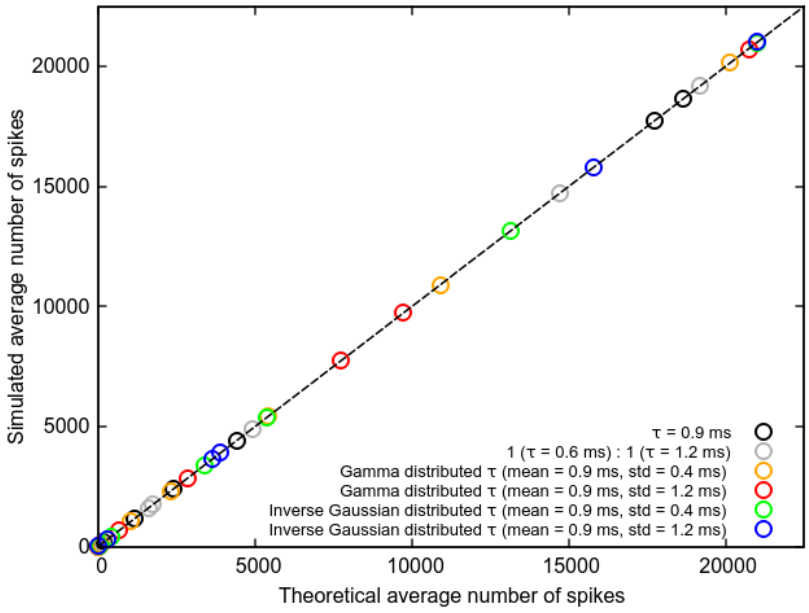

Figure S1: Comparison between the simulated and theoretical average numbers of spikes in a time scan predicted by Eq. 4 (see article) for different distributions (constant-valued, bimodal, Gamma and inverse Gaussian) of the particle event duration τ. All data points have been obtained from 10,000 Monte Carlo runs with $\langle\tau\rangle = 900$ μs. For the Gamma and the inverse Gaussian distributions, two values for the standard deviation of τ have been used: 400 and 1200 μs. For the bimodal distribution, half of the particle events were supposed to last 600 μs and half 1200 μs. The dashed line is here to guide the eye.

S3 Normality of the spike count

Fig. S2 displays the histogram of spike counts obtained from 5,000 simulated time scans with $\lambda = 10$, 100 and 1000 s^{-1}. Visually, these histograms look Gaussian. The Shapiro-Wilk test confirms that these histograms can indeed be considered as stemming from a normal probability distribution. The values W of the test statistic and the corresponding p-values are given in Tab S2.

Table S2: Shapiro-Wilk test statistic W values and corresponding p-values for the three simulated nanoparticle flux rates.

	W test statistic	p-value
$\lambda = 10\,\mathrm{s}^{-1}$	0.99953	0.27
$\lambda = 100\,\mathrm{s}^{-1}$	0.99949	0.19
$\lambda = 1000\,\mathrm{s}^{-1}$	0.99959	0.39

S4 Operating parameters

Pneumatic nebulizer
Solution uptake rate: 0.3 rps ($\approx 0.346\,\mathrm{mL\,min}^{-1}$)

Torch
Material: quartz
Internal diameter: 1.5 mm

Plasma
RF power: 1550 W
Sampling depth: 7 mm
Carrier gas: $0.84\,\mathrm{L\,min^{-1}}$ (for measurements with $\tau_{\mathrm{dw}} = 100\,\mu$s) and $0.74\,\mathrm{L\,min^{-1}}$ (for measurement with $\tau_{\mathrm{dw}} = 300\,\mu$s)

Electron multiplier
Dead time: 40 ns

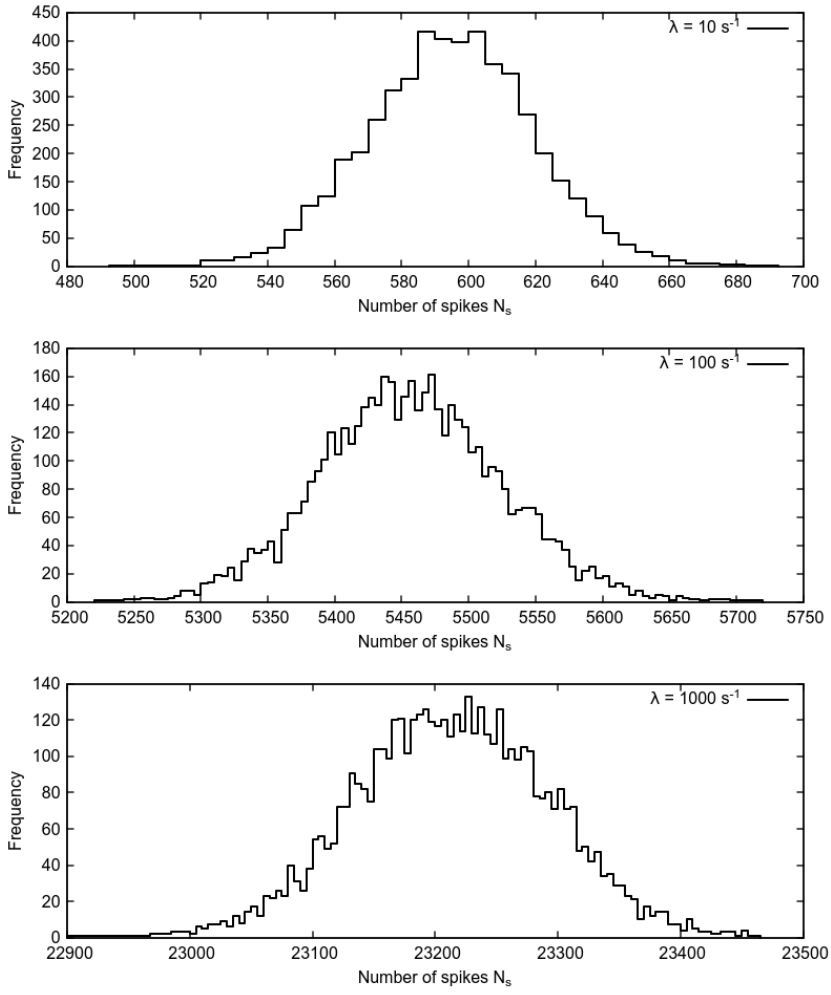

Figure S2: Empirical distribution of the number of spikes N_s for three different nanoparticle flux rates (top: $\lambda = 10\,\mathrm{s^{-1}}$; middle: $\lambda = 100\,\mathrm{s^{-1}}$; bottom: $\lambda = 1000\,\mathrm{s^{-1}}$) obtained by Monte Carlo simulation (5,000 runs), for a 60 s experiment with a 100 μs dwell time and $\langle\tau\rangle = 800\,\mu$s. The bin width of the histograms was set equal to 5.

S5 Comparison between spike detection algorithms

Tab. S3 provides a comparison between the number of spikes detected with our algorithm and with MassHunter, Agilent's proprietary software. Very different values are obtained. The details of the algorithm used in MassHunter are not publicly available, but the general approach is known. Once the baseline has been found by iteration, spikes present in the signal are integrated and counted. This

method is known to work well to determine the particle number concentration of dilute nanoparticle dispersions. By contrast, in the present study, a spike is simply defined as a contiguous sequence of nonzero readings in a time scan and detected accordingly. Consequently, depending on the magnitude of the nanoparticle flux rate, a spike can display only one or many local maxima and some of them may be considered as distinct spikes by MassHunter. In the highly concentrated regime, MassHunter is indeed impacted by the existence of a nonzero "effective background" that the software interprets as a dissolved background, whereas it stems from the overlap of many particle events. It explains why the number of spikes found by MassHunter can be much higher than the number of spikes detected by our algorithm for highly concentrated dispersions. We emphasize again that our definition of a spike, despite its sheer simplicity, differs from the one classically employed and implemented in MassHunter for spike detection, and is crucial for the accordance with the theory.

Table S3: Comparison between the mean spike counts detected with our own custom-written program and with MassHunter.

Dilution factor	Mean spike count (custom-written program)	Mean spike count (MassHunter)
250000	999	999
7500	20218	14227
3750	23142	17102
2500	19164	34093
1875	14535	6695
1500	10231	6817
1250	6736	6801
1071	4187	6769
938	2873	6791
833	1865	6725
750	1053	6719
682	608	6624
625	395	6615
577	258	6715
536	134	6566
500	95	6619
250	1	6068

S6 Single particle ICP-MS analysis of dispersions composed of a mixture of 60 and 150 nm-sized nanoparticles

In order to check that the theory remains valid when a nanoparticle dispersion is made up of particles of different sizes, we performed some sp-ICP-MS experiments with dispersions composed of a mixture of 60 and 150 nm-sized nanoparticles. 60 and 150 nm gold nanoparticle dispersions stabilized with a citrate buffer have been purchased from Sigma-Aldrich. The particle number concentration of the 60 nm stock dispersion was $\approx 1.9 \times 10^{10}\,\mathrm{mL}^{-1}$, while it was $\approx 3.6 \times 10^{9}\,\mathrm{mL}^{-1}$ for the 150 nm stock dispersion.

The most dilute dispersion we analyzed was made from a mixture between the $250,000\times$ diluted 60 nm dispersion and the $150,000\times$ diluted 150 nm dispersion. According to the particle number concentrations of the stock solutions, 76% of the particles of this dispersion had a 60 nm size and 24% a 150 nm size. The proportion of 150 nm particles found by sp-ICP-MS, $(24 \pm 2)\%$, is consistent with this estimation. We also analyzed more concentrated mixed dispersions, always with the same ratio between the numbers of 60 and 150 nm sized nanoparticles.

Following the procedure described in the "Experimental validation" section of the article, we got the results plotted in Fig S3. The continuous curve was obtained thanks to the method presented in the subsection 3.4 ("Comparison with theory") and fits the data points very well; the average duration of

particle events $\langle\tau\rangle$ was found equal to 792.5 μs. These experimental results confirm that the theory works just as well for dispersions composed of differently sized nanoparticles as it does for single-sized nanoparticles.

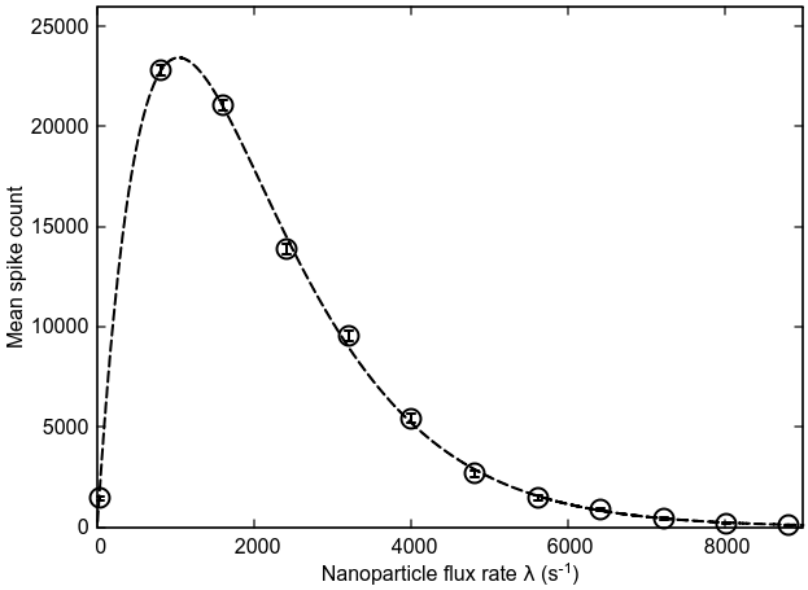

Figure S3: Mean spike counts extracted from the experimental time scans of 3:1 mixtures of 60 and 150 nm-sized nanoparticles, for $\tau_{\mathrm{dw}} = 100\,\mu$s. The dashed line is the numerical adjustment to the model ($\langle\tau\rangle = 792.5\,\mu$s). The three-sigma error bars associated with each data point were computed by Monte Carlo simulation.