



HAL
open science

Analysis of gamma-ray spectra with spectral unmixing -Part I: Determination of the characteristic limits (decision threshold and statistical uncertainty)

Jiaxin Xu, Jérôme Bobin, Anne de Vismes Ott, Christophe Bobin, Paul
Malfrat

► To cite this version:

Jiaxin Xu, Jérôme Bobin, Anne de Vismes Ott, Christophe Bobin, Paul Malfrat. Analysis of gamma-ray spectra with spectral unmixing -Part I: Determination of the characteristic limits (decision threshold and statistical uncertainty). 2021. hal-03345737

HAL Id: hal-03345737

<https://hal.science/hal-03345737>

Preprint submitted on 15 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analysis of gamma-ray spectra with spectral unmixing - Part I: Determination of the characteristic limits (decision threshold and statistical uncertainty)

Jiaxin Xu^a, Jérôme Bobin^{b,*}, Anne de Vismes Ott^a, Christophe Bobin^c, Paul Malfrait^a

^a*Institut de Radioprotection et de Sécurité Nucléaire (IRSN), PSE-ENV, SAME, LMRE, Orsay, 91400, France*

^b*IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*

^c*Université Paris-Saclay, CEA, List, Laboratoire National Henri Becquerel (LNE-LNHB), F-91120 Palaiseau, France*

Abstract

Poisson-statistics based spectral unmixing has been shown to be an efficient analysis tool for the radionuclides activity estimation from gamma-ray spectrometry measurements. However, the calculation of the corresponding characteristic limits has not been investigated so far. In this paper, we present the quantification of the decision threshold and the limits of the coverage interval for the metrological use of such spectral unmixing algorithms. The proposed approach is evaluated and validated with simulated spectra of HPGe and NaI measurements by comparing the results to characteristic values calculated from Monte Carlo simulations. We focus particularly on the validation of the method for the metrological analysis of environmental measurements, for which the low-level activity quantification requires an accurate characteristic limits determination. Along with the instrument calibration studied in Xu et al. (2021), we establish a metrological analysis tool by using the spectral unmixing algorithm for environmental aerosol gamma-ray measurements.

Keywords: gamma-ray spectrometry, decision threshold, uncertainty, full spectrum analysis, spectral unmixing

*Corresponding author

Email address: `jerome.bobin@cea.fr` (Jérôme Bobin)

1. Introduction

Gamma-ray spectrometry is one of the main techniques used for the identification and quantification of gamma-emitting radionuclides. The measurements can be performed with different types of detectors, including semi-conductor detectors such as High Purity Germanium detectors (HPGe) and the scintillation detectors, *e.g.*, Sodium Iodide (thallium doped), NaI(Tl). A gamma-ray spectrum is the histogram of the number of detected events as a function of the energy that is deposited by the gamma-ray or the X-ray in the detector. In this context, radionuclides are classically identified and quantified with peak-based or deconvolution-based spectrum analysis methods (*e.g.*, Mirion Technologies (Canberra)'s Genie 2000 gamma analysis software ¹), which however has limited performances in terms of low-level activity measurements. In recent studies, with the purpose of providing more accurate and sensitive activity measurements, spectral unmixing methods have been investigated for gamma-ray spectrum analysis in Xu et al. (2020) applied to HPGe measurements and in Paradis et al. (2020), André et al. (2020) applied to NaI measurements.

In contrast to peak-based analysis only from the full energy peaks, spectral unmixing is an analysis tool that identifies and quantifies radionuclides from the full energy spectrum (*i.e.*, peaks and their associated continua) of a gamma-ray measurement. Furthermore, it allows accounting for the Poisson statistics that underlie the detection process. In this framework, gamma-ray spectrum analysis is recast as a regularized inverse problem, with the aim of decomposing a measured spectrum into the individual spectral contribution of each radionuclide present in the measured sample. A measured spectrum \mathbf{x} is composed of M channels: $\mathbf{x} = [x_1, \dots, x_M]$. For $\forall i \in [1, \dots, M]$, the Poisson process of radioactive decay leads to model the problem as:

$$x_i \sim \text{Poisson}([\Phi \mathbf{a}]_i + b_i) \quad (1)$$

¹<https://www.mirion.com/products/genie-2000-gamma-analysis-software>

where we note the radionuclides' spectral signatures as $\Phi = [\phi_1, \dots, \phi_N]$ (containing N radionuclides' spectral signatures) and the background spectrum as \mathbf{b} . The spectral unmixing addresses the estimation of radionuclides' mixing weights vector \mathbf{a} that are proportional to radionuclides' activities. According to the Poisson statistics, the likelihood of measuring a spectrum \mathbf{x} knowing that the spectra linear combination is $\Phi\mathbf{a} + \mathbf{b}$ can be written as follows:

$$\mathcal{P}(\mathbf{X} = \mathbf{x} | \Phi\mathbf{a} + \mathbf{b}) = \prod_i^M \frac{[\Phi\mathbf{a} + \mathbf{b}]_i^{x_i} e^{-[\Phi\mathbf{a} + \mathbf{b}]_i}}{x_i!} \quad (2)$$

Minimizing the neg-log-likelihood of this distribution leads to the following estimator of the mixing weights (Xu et al., 2020):

$$\hat{\mathbf{a}} = \underset{\mathbf{a} \geq 0}{\operatorname{argmin}} \Phi\mathbf{a} + \mathbf{b} - \mathbf{x} \odot \log(\Phi\mathbf{a} + \mathbf{b}) \quad (3)$$

where \odot is the Hadamard product and $\hat{\mathbf{a}}$ (*i.e.* entry-wise product) represents the estimate of the mixing weights vector \mathbf{a} . The above optimization problem has been tackled as a regularized inverse problem, where the non-negativity is imposed on the solution to be estimated.

The Poisson-statistics based spectral unmixing has been shown to be a more accurate activity estimation method than peak-based analysis. To use such estimation algorithms for experimental data analysis, quantifying uncertainties is a critical step for assessing the statistical confidence of the estimation procedure as well as for decision making purposes. In radioactivity metrology, it is conventionally called the characteristic limits of a measurement. These limits comprise the decision threshold, the detection limit and the limits of the coverage interval, they are essential to achieve a metrological use of the spectral unmixing procedure. The determination of these limits are defined in the ISO 11929 (2010). Their assessment has been studied for commonly used peak-based gamma-ray spectrum analysis method. Korun et al. (2014) studies the calculation of the decision thresholds. Boshkova (2018) investigates the statistical significance of the measured net signal through the determination of characteristic limits. So far, and to the best of our knowledge, no study has considered the full spectrum analysis case. The aim of this work is to explore how the evaluation of the

characteristic limits can be generalized to the Poisson-statistics based spectral unmixing procedure.

The paper is organized as follows: Section 2 presents the concept of characteristic limits for radioactivity measurements. In Section 3, we describe the data used in this work. Section 4 explores the calculation of the decision threshold. Next, we discuss the assessment of the coverage intervals in Section 5. Section 7 provides the conclusions of this work.

2. Characteristic limits in radioactivity measurements

2.1. Definitions

Before going further, we first introduce some useful notations to describe the characteristic limits:

- A : Measurand, the quantity of interest.
- \hat{a} : Determined value of the measurand Y (*i.e.*, the estimate of A).
- \tilde{a} : True value of the measurand.

[JB: Pourquoi ne pas mettre les mixing weights en gras, ils le sont plus loin ?]
JX: dans ce contexte, il s'agit l'estimation de l'activity d'un radionuclide, *i.e.*, mixing weigt, donc c'est un scalar

Quantifying the characteristic limits in gamma-ray spectrum analysis is carried out within the classical statistical hypothesis framework. It considers testing hypotheses with the two following alternatives (associated with type I error and type II error described in Table 1):

- H_0 : the null hypothesis, where a given radionuclide is not “active”, *i.e.*, namely not present in the sample.
- H_1 : the alternative hypothesis, where the radionuclide is present in the mixture.

The standardization document (ISO 11929, 2010) defines the characteristic limits for ionizing radiation measurements and provides a framework for the computation of these limits. Referring to Weise et al. (2005), Michel (2016), the definition and interpretation of these limits for some estimate \hat{a} of a measurand A are as follows:

- **Decision threshold (DT)** allows a decision to be made on whether or not the physical effect quantified by the measurand is present.

The determination of DT is related to the Type I error described in Table 1. When the quantity y exceeds the critical value (DT), the null hypothesis H_0 should be rejected with respect to a given false positive rate (FPR). It can be described with:

$$\alpha = \mathcal{P}(\hat{a} \geq DT | \tilde{a} = 0) \quad (4)$$

where \tilde{a} is the true value of the measurand and α is the desired critical FPR.

- **Detection limit (DL)** indicates the smallest true quantity value of the measurand, which can still be detected with the applied measurement procedure.

The determination of DL is related to the Type II error described in Table 1. It is selected with respect to a desired false negative rate (FNR) based on the decision threshold level.

More precisely, the detection limit (DL) is the smallest value that provides a desired Type II error probability β :

$$\beta = \mathcal{P}(\hat{a} \leq DT | \tilde{a} = DL) \quad (5)$$

where the DT is given and \tilde{a} is the true value of the measurand.

- **The coverage interval** for the estimate \hat{a} is an interval that has a probability $1 - \gamma$ of containing the true value \tilde{a} :

$$1 - \gamma = \mathcal{P}(\hat{a}_l < \tilde{a} < \hat{a}_r) \quad (6)$$

where $[\hat{a}_l, \hat{a}_r]$ stands for the coverage interval.

The regular radioactivity monitoring throughout France is ensured by the operators, regulated by the ASN (French Nuclear Safety Authority) and supplemented by IRSN (French Radiation Protection and Nuclear Safety Institute). The environmental samples measurements are performed in this framework and the measurements data are fed into the RNM network ¹. For that purpose ASN imposes that data contain the activity, the expanded uncertainty and the decision threshold. Indeed the decision threshold serves to conclude if the measurement is statistically significant or not, given a fixed false positive rate α . The detection limit is not required because it is supposed to be compared with some regulatory levels, while the radioactivity monitoring aims to detect rapidly any increase in environmental radioactivity.

In this paper, we thus focus on the determination of the decision threshold and the coverage interval for Poisson-statistics based spectral unmixing algorithm. Firstly, one can determine whether the resulting activity of a radionuclide is significant by comparing it to the decision threshold. Secondly, the coverage interval provides the statistical uncertainty of the estimation.

2.2. Quantification of characteristic limits with Monte Carlo simulations

A traditional approach to quantify the characteristic limits for some estimation method is to make use of Monte Carlo simulations. The main drawback is their massive computational cost since Monte-Carlo simulations are needed for each new spectrum to be analyzed. Furthermore, the true value of the measurement in a measured spectrum is not known in advance. Therefore, we focus on alternatives which are less computationally demanding and yet precise to derive the characteristic limits without resorting to Monte-Carlo simulations. To evaluate and validate these methods, we propose to use synthetic spectra in which the radionuclides' mixtures are perfectly known. This approach was employed because the characteristic limits of these synthetic mixtures can be precisely

¹<https://www.mesure-radioactivite.fr/en#/expert>

calculated by Monte Carlo simulations, which allows assessing the statistical uncertainty of the proposed approach.

Decision threshold from Monte Carlo simulations. Let the probability distribution of the mixing weight under null hypothesis of the radionuclide indexed by j be defined as:

$$\mathcal{P}(a_j | \tilde{a}_j = 0) \tag{7}$$

According to Eq.(4), the decision threshold can be derived from $1 - \alpha$ percentiles of this probability distribution, for α being the false positive rate with respect to the DT level. In practice, $\alpha = 2.5\%$ is usually used in environmental measurements. Assuming that the linear combination model of radionuclides is known, one can make use of Monte Carlo simulations that mimic the mixture under the null hypothesis of each radionuclide to evaluate their DT levels.

In the experiments of this paper, 1000 simulations are performed to calculate an accurate DT value according to $\alpha = 2.5\%$ for each radionuclide. In each random simulation process, we set the radionuclide to zero and other radionuclides to their actual levels. The spectral unmixing algorithm is subsequently applied to estimate radionuclides' mixing weights from these simulated spectra. The results obtained for the radionuclide set to zero (*i.e.*, null hypothesis) allow us to compute its decision threshold value.

Coverage interval from Monte Carlo simulations. Since each estimate of the mixing weights is assumed to have a limited estimation bias or error, a coverage interval with respect to a given probability γ for individual radionuclides can also be derived by performing Monte-Carlo simulations. In practice, $\gamma = 5\%$ is usually taken into account. More precisely, we first perform 1000 Monte Carlo simulations that mimic the actual mixing scenario of radionuclides. Next, the spectral unmixing algorithm is applied to estimate mixing weights from these simulated spectra. Finally, the coverage intervals of radionuclides can be derived from the distribution of their estimated mixing weights.

3. Data description

The main contribution in this paper is the characteristic limits calculation with spectral unmixing analysis, applied to aerosol measurements performed with HPGe detectors. The validated methods are further studied in Xu et al. (2021) for quantitative analysis of experimental data. In this context, we make use of simulations of aerosol filter measurements of a Mirion Broad Energy Germanium (BEGeTM) (61 % relative efficiency) investigated in Xu et al. (2020), covering the energy range from 20 keV to 1640 keV composed in 16384 channels. The choice of gamma-ray emitters consists of 10 radionuclides: ⁷Be, ²²Na, ⁴⁰K, ¹³⁷Cs, ²¹⁰Pb, ²⁰⁸Tl, ²¹²Bi, ²¹²Pb, ²¹⁴Bi, ²¹⁴Pb. The background radiation spectrum is according to a measurement without sample during 560000 s and normalized to the simulation counting time. The activity of the blank filters is disregarded since it has been checked experimentally to be negligible. Table 2 shows the activities of these radionuclides that are customary in real aerosol measurements (*i.e.*, measured during 320000 s). Since each simulated spectral signature corresponds to the energy response with unit particle (*i.e.* one disintegration), the relationship of the radionuclides' mixing weights vector \mathbf{a} and their activities can be described as follows (see details in Xu et al. (2021)):

$$\mathbf{g} = \frac{\mathbf{a}}{t} \quad (8)$$

where \mathbf{g} is the radionuclides' activities in becquerel (Bq) (*i.e.*, number of disintegrations per second) and t is the counting time.

We focus on the assessment of characteristic limits for 4 radionuclides: ⁷Be, ²²Na, ¹³⁷Cs, ²¹²Pb, since these radionuclides cover the whole energy range and different statistical regimes. The simulation model of 10 radionuclides and the contributions of these 4 radionuclides are displayed in Fig. 1.

Next, we also generate simulations of measurements of a 3"x3" NaI(Tl) detector without shielding using point sources placed at a distance of 1 m (Paradis et al., 2020), which is made of 1024 channels. The spectral signatures corre-

spond to the detector response of 4 gamma-emitting radionuclides with photon emissions covering a range of energies between 40 keV and 2 MeV: ^{60}Co , ^{134}Cs , ^{137}Cs , ^{152}Eu (see Fig. 2). We make use of simulated spectra with respect to the numbers of counts shown in Table 3.

4. Quantifying the decision threshold

In peak-based gamma-ray spectrum analysis, the decision threshold is usually derived from some statistical test based on the measured spectrum. This means to evaluating how much the estimated quantity associated with a radionuclide's activity departs from the background (*i.e.*, other contributions composed in the measured spectrum) and is therefore statistically consistent or not with this background.

In the full spectrum analysis context, the spectral unmixing decomposes a gamma-ray spectrum into individual spectra of radionuclides. To determine the decision threshold of a single radionuclide indexed by j in the unmixing model, we reformulate the true linear mixing model with this radionuclide and an equivalent background:

$$\Phi \mathbf{a} + \mathbf{b} \rightarrow \phi_j a_j + \mathbf{m} \quad (9)$$

where $\phi_j a_j$ represents the contribution of the j^{th} radionuclide in the spectrum. The equivalent background is then composed of the background spectrum \mathbf{b} and the contribution of all but the tested radionuclide:

$$\mathbf{m} = \sum_{l \neq j} \phi_l a_l + \mathbf{b} \quad (10)$$

Take an example of ^{137}Cs in the measurement presented in Fig. 1, the spectral contributions are illustrated in Fig. 3.

Recall the definition in Section 2, the decision threshold of the measurand a_j (*i.e.*, the mixing weight of the j^{th} radionuclide) is derived from:

$$\alpha = \mathcal{P}(a_j \geq DT | \tilde{a}_j = 0) \quad (11)$$

The null hypothesis of \tilde{a}_j (*i.e.*, true value of a_j) implies that \mathbf{m} is the mean value of the distribution:

$$\mathbf{x} \sim \text{Poisson}(\mathbf{m}) \quad (12)$$

We assume that \mathbf{m} is accurately estimated by the spectral unmixing method. This estimate is denoted by $\hat{\mathbf{m}}$. The decision threshold of the j^{th} radionuclide can then be derived by the following statistical test:

$$\alpha = \mathcal{P}(t \geq T(\boldsymbol{\lambda}) | \lambda_i = \hat{m}_i, \forall i \in [1, \dots, M]) \quad (13)$$

where M is the number of channels in the measured spectrum. The quantity $T(\boldsymbol{\lambda})$ is some test statistics that is defined as a scalar function of some background model $\boldsymbol{\lambda}$. The decision threshold can be derived from the statistical distribution of such a test statistics under the null hypothesis: the tested radionuclide is not present and the observed spectrum is described by the equivalent background only. Next, we will consider different choices for the $\boldsymbol{\lambda}$ and $T(\boldsymbol{\lambda})$. More precisely, we will consider test statistics $T(\boldsymbol{\lambda})$ that will be defined of weighted sums of the total number of counts:

$$T(\boldsymbol{\lambda}) = \sum_{i=1}^M w_i x_i \quad (14)$$

Thanks to the statistical independence of each channel of the measured spectrum, the distribution of such general test statistics can be written by the following generic expression:

$$\mathcal{P}\left(\sum_{i=1}^M w_i x_i \mid \sum_{i=1}^M w_i \hat{m}_i\right) \quad (15)$$

where $\mathbf{w} = [w_1, \dots, w_M]$ stands for a vector of weights. The goal of this general approach is that it allows to select weights that are better adapted to distinguish between the radionuclides' contribution in the measured spectrum and its equivalent background. The DT (noted a_j^*) of the j^{th} radionuclide with respect to a given false positive rate α , are studied with 4 different choices of \mathbf{w} .

1. A simple statistical test to consider is based on the total number of counts, as measured by the sum of observed counts under the null hypothesis $\tilde{a}_j = 0$, which leads to:

$$\mathbf{w}_{sum} = \{w_i = 1, \forall i \in [1, \dots, M]\} \quad (16)$$

Eq.(15) can be formulated with the following Poisson distribution:

$$\sum_{i=1}^M x_i \sim \mathcal{Poisson} \left(\sum_{i=1}^M \hat{m}_i \right) \quad (17)$$

The DT, a_j^* can be derived from its cumulative distribution function (CDF):

$$\alpha = \mathcal{P} \left(\sum_{i=1}^M x_i \geq \sum_{i=1}^M [\phi_j a_j^*]_i + \sum_{i=1}^M \hat{m}_i \right) \quad (18)$$

2. By using \mathbf{w}_{sum} , we take into account the information carried out by the full spectrum, it however poorly distinguishes the radionuclide to be tested from the background. For instance, in Fig. 3, the Compton continuum of the ^{137}Cs is under its equivalent background, while it is better distinguished from the background in the peak region. Therefore, we rather use the pre-specified channels in a region of interest (ROI - *e.g.* peak region of the radionuclide), which can be written as:

$$\mathbf{w}_{ROI} = \begin{cases} w_i = 1, & \text{if } i \in ROI \\ w_i = 0, & \text{otherwise} \end{cases} \quad (19)$$

Indeed, this approach is carried out in the same manner as peak-based analysis, in which the DT level is derived based on the number of background counts (ISO 11929, 2010) for more details).

We calculate the DT from the CDF of the following distribution:

$$\sum_{i \in ROI} x_i \sim \mathcal{Poisson} \left(\sum_{i \in ROI} \hat{m}_i \right) \quad (20)$$

3. In order to better distinguish between the radionuclide to be tested and its equivalent background in the full energy range, we further investigate

the DT calculation based on the measured spectrum \mathbf{x} weighted by the projection of \mathbf{x} onto ϕ_j parallelly to \mathbf{m} , which is graphically illustrated in Fig. 4. In contrast to a standard orthogonal projection, a projection onto ϕ_j parallelly to \mathbf{m} allows to measure the contribution of \mathbf{x} explained by the signature ϕ_j while guaranteeing its contribution along the direction \mathbf{m} to be zero. This eventually enhances the contrast between the contribution of the j -th radionuclide and the equivalent background \mathbf{m} in \mathbf{x} , thus providing a more sensitive evaluation of the detection threshold. Such a parallel projection is denoted by \mathbf{w}_{pb} .

In practice it can be evaluated by first defining the 2-columns matrix $\Psi = \begin{bmatrix} \phi_j & \hat{\mathbf{m}} \end{bmatrix}$: the spectral signature ϕ_j and the estimated equivalent background $\hat{\mathbf{m}}$. A simple weights vector to be considered is:

$$\mathbf{w}_{pb} = \{(\Psi^T \Psi)^{-1} \Psi^T\}_{\phi_j} \quad (21)$$

where $\{\}_{\phi_j}$ defines the entry related to ϕ_j .

However, the resulting statistical test does not follow a Poisson distribution. Thankfully, since it is defined as a linear combination of a large number of observed channels, one can approximate the distribution to be Gaussian. The resulting statistical test with observed counts \mathbf{x} and estimated equivalent background \mathbf{m} will be calculated as follows:

$$\sum_{i=1}^M [w_{pb}]_i x_i \sim \mathcal{N} \left(\sum_{i=1}^M [w_{pb}]_i \hat{m}_i, \sum_{i=1}^M [w_{pb}]_i^2 \hat{m}_i \right) \quad (22)$$

while in each channel, the mean value $\mu = [w_{pb}]_i \hat{m}_i$ and variance $\sigma^2 = [w_{pb}]_i^2 \hat{m}_i$ are considered due to the Poisson statistics of the model.

The DT, a_j^* can be derived from the CDF of the above distribution:

$$\frac{\alpha}{2} = \mathcal{P} \left(\sum_{i=1}^M [w_{pb}]_i x_i \geq \sum_{i=1}^M [w_{pb}]_i [\phi_j a_j^*]_i + \sum_{i=1}^M [w_{pb}]_i \hat{m}_i \right) \quad (23)$$

4. The above statistical test assumes the underlying noise is additive, white and Gaussian. To consider the covariance matrix underlying the Poisson noise, we further make use of an approach based on the a re-weighted

projection parallel to the background, noted \mathbf{w}_{rpb} :

$$\mathbf{w}_{rpb} = \left\{ \left(\Psi^T \text{diag} \left(\frac{1}{\phi_j \hat{a}_j + \hat{\mathbf{m}}} \right) \Psi \right)^{-1} \Psi^T \text{diag} \left(\frac{1}{\phi_j \hat{a}_j + \hat{\mathbf{m}}} \right) \right\}_{\phi_j} \quad (24)$$

and calculate the DT from the CDF of:

$$\sum_{i=1}^M [w_{rpb}]_i x_i \sim \mathcal{N} \left(\sum_{i=1}^M [w_{rpb}]_i \hat{m}_i, \sum_{i=1}^M [w_{rpb}]_i^2 \hat{m}_i \right) \quad (25)$$

Experimental results

These four proposed statistical tests are applied to assess the decision threshold for HPGe and NaI measurements described in Section 3. The results are compared to those carried out with Monte Carlo simulations, noted as MC (described in Section 2.2). The false positive rate is fixed to $\alpha = 2.5\%$. More precisely, the proposed statistical tests are used to assess the decision threshold of radionuclides in the simulations of the HPGe and NaI measurements for each Monte Carlo simulation. The statistical test with \mathbf{w}_{ROI} is only used for HPGe measurements. Indeed, NaI measurements have much lower spectral resolution, which makes it much harder to define an accurate ROI. Experiments on the described selection of gamma-ray emitters (see Fig. 1) of the HPGe measurement considers the following peak regions: ${}^7\text{Be}$ at 477 keV, ${}^{22}\text{Na}$ at 1274 keV, ${}^{137}\text{Cs}$ at 662 keV and ${}^{212}\text{Pb}$ at 238 keV.

For HPGe spectra, the different statistical tests are compared in Fig. 5, 6, 7 and 8 for ${}^7\text{Be}$, ${}^{22}\text{Na}$, ${}^{137}\text{Cs}$ and ${}^{212}\text{Pb}$ respectively. For NaI spectra, the statistical tests are compared in Fig. 9, 10, and 11, and 12 for ${}^{60}\text{Co}$, ${}^{134}\text{Cs}$, ${}^{137}\text{Cs}$ and ${}^{152}\text{Eu}$ respectively. The results show the distributions of:

- the accurate DT level calculated from Monte Carlo (*i.e.*, $1 - \alpha$ percentile of the distribution under the null hypothesis) is displayed with a dotted line (red).
- Distribution of calculated DT based on the statistical test of equivalent background weighted by \mathbf{w}_{sum} (red), \mathbf{w}_{ROI} (green), \mathbf{w}_{pb} (blue) and \mathbf{w}_{rpb} (orange).

From these results, we can draw the following conclusions:

- Firstly, the statistical test based on the equivalent background weighted by \mathbf{w}_{sum} is less precise to derive the DT than other approaches, since it poorly distinguishes the radionuclide’s spectrum from the background.
- Compared to statistical tests based on the projected equivalent background (*i.e.*, weighted by \mathbf{w}_{pb} and \mathbf{w}_{rpb}), the ROI-based statistical test (*i.e.*, weighted by \mathbf{w}_{ROI}) is less precise since it only uses information at the vicinity of the peak regions.
- For HPGe measurement, for ^{22}Na , ^{137}Cs and ^{212}Pb , statistical tests based on the equivalent background weighted by \mathbf{w}_{rpb} derives similar DT levels compared to the actual $1 - \alpha$ percentile of the distribution (*i.e.*, accurate DT level), which is better than the choice of w_{pb} . This is not valid for ^7Be due to the fact that ^7Be is predominant in the measured spectrum, while its equivalent background is poorly estimated. However, the determination of an accurate DT level for low-level radionuclides is more important for decision making. Indeed, in such a case, the dominant spectral contribution, such as ^7Be , generally yield a more precise estimation of the equivalent background, which entail a more accurate DT determination for low-level radionuclides.
- When we further focus on the results of NaI measurements, the choice of using w_{rpb} is shown to be more consistent than w_{pb} . It allows to better distinguish a radionuclide from its equivalent background when the spectra have a significant overlap in the whole energy range.

The DT assessment is evaluated with Monte Carlo simulations of realistic measurements of aerosol samples performed with HPGe detectors. The proposed analysis involving a statistical test based on re-weighted projections of the equivalent background is shown to be an efficient and rapid tool to determine DT. We apply this procedure to real measurements in the companion article Xu et al. (2021).

5. Coverage interval

Evaluating the statistical reliability of a parameter estimation procedure is generally composed of two elements: the estimation bias and an interval of values that contains the true value of the parameter up to some probability. In the spectral unmixing approach, the statistical uncertainty assessment can be carried out from Monte Carlo simulations that mimic the actual mixture of radionuclides as presented in Section 2.2. The distribution of estimated values provides both the estimation bias and the coverage interval. However, this requires a massive amount of simulations, making this approach hard to implement in practice. In this study, we focus on the metrological use of the Poisson-based spectral unmixing for analyzing the described HPGe detectors performed aerosol measurements. In Xu et al. (2020), we showed that estimation biases with spectral unmixing analysis are low for routine aerosol measurements.

In this context, we propose to make use of the Fisher information matrix of the estimated mixing weights to compute the coverage interval. This allows for an analytic expression of the coverage interval, which does not require computationally expensive Monte Carlo simulations. Such method has been also investigated in André et al. (2020) for the coverage intervals assessment of spectral unmixing approach applied to NaI measurements.

For the observed variable x distributed as $f(x|\theta)$, the asymptotic distribution of the maximum likelihood estimator (MLE) of the parameter θ is a Gaussian distribution:

$$\mathcal{N}(\theta, I(\theta)^{-1}) \quad (26)$$

where $I(\theta)$ is the Fisher information (Fisher, 1956) defined as:

$$I(\theta) = \mathbb{E}_\theta \left[\frac{\partial^2 \log f(x|\theta)}{\partial \theta^2} \right] \quad (27)$$

where \mathbb{E}_θ stands for the expectation taken with respect to $f(x|\theta)$. The true value of the parameter is generally not known. When the estimation bias $|\hat{\theta} - \theta|$ is assumed to be low, the estimation of standard error of θ_{MLE} can be obtained by replacing the unknown true value θ by the estimated value $\hat{\theta}$. To validate

the metrological use of such an approach, we will compare the coverage interval obtained from the information matrix associated with $\hat{\theta}$ to the one we would obtain from Monte Carlo simulations.

Recall that the Poisson-based spectral unmixing provides a maximum likelihood estimate of the mixing weights, noted $\hat{\mathbf{a}}$. We propose a first approximation by considering that the estimator of the mixing weights vector is non-biased and follows a Gaussian distribution centered at $\hat{\mathbf{a}}$. According to the Poisson likelihood Eq.(3), the covariance matrix is approximated with the inverse of the Fisher information matrix:

$$I(\hat{\mathbf{a}})^{-1} = \left(\Phi^T \text{diag}(\mathbf{x} \oslash (\Phi \hat{\mathbf{a}} + \mathbf{b})^2) \Phi \right)^{-1} \quad (28)$$

where \oslash is the element-wise division operator. The coverage interval of the estimated mixing weights $\hat{\mathbf{a}}$ can be derived from the diagonal elements of the inverse of the Fisher matrix; the resulting standard deviation for $\hat{\mathbf{a}}_i$ is then defined as $\sqrt{[I(\hat{\mathbf{a}})^{-1}]_{i,i}}$, which approximates the standard deviation of the maximum likelihood estimator distribution. This approximation will be adapted when estimated activity is high enough so that the coverage interval does not contain 0. In Section 6, we discuss how the proposed approach can be extended to this case.

Experimental results

The coverage intervals assessment with Fisher information matrix is evaluated with the simulations of HPGe and NaI measurements described in Section 3. To validate this approach, the Gaussian distribution approximated with Fisher information matrix associated with the estimated mixing weights need to agree with the expected distribution (*i.e.*, estimated distribution obtained from Monte Carlo simulations). Firstly, we make use of the Q-Q (quantile-quantile) plots of the Normal distribution generated from the standard deviation calculated with the Fisher information and the distribution of estimations derived from the Monte Carlo simulations, by plotting their quantiles one against another. The aim is to i), test the normality of the estimator, which allows val-

idating the coverage interval with a standard uncertainty in terms of Normal distribution. ii), compare the standard deviation of each estimation calculated with the Fisher information to those obtained with Monte Carlo simulations. More precisely, the Q-Q plots are shown in Fig. 13 and Fig. 15 for the HPGe measurement and the NaI measurement, respectively.

The Q-Q plots allows comparing the distribution of the scales and centered estimated parameters with a centered normal distributed, which is its expected asymptotic distribution. The straight lines show that taling the asymptotic Normal approximation provides a very accurate evaluation of the distribution of the estimated mixing weights, at least in the interval $[0, 3]$. [JB: taling dans ce context ?] Additionally, we can asses how accurate are the coveral intervals derived form the Fisher information matrix. For that purpose, we perform Monte-Carlo simulations to compute the probability associated with the intervals : $a_0 \pm \sigma_f$, $a_0 \pm 2\sigma_f$ and $a_0 \pm 3\sigma_f$, where a_0 , σ_f stand for the expected mixing weight (*i.e.*, simulated value), which we comapre with the error determined by the Fisher matrix. The corresponding percentages are expected to be 68.27 %, 95.45 % and 99.73 % for Normal distribution. The results are shown in Table 4 and Table 5 for HPGe and NaI measurements respectively, which confirms that for this measurement scenario, the coverage intervals computed from the Fisher information matrix are quite accurate.

The results in Table 4 show that the quantiles of the estimation distribution of Monte Carlo simulations within: $a_0 \pm \sigma_f$, $a_0 \pm 2\sigma_f$ and $a_0 \pm 3\sigma_f$ are consistent with their expected values for ${}^7\text{Be}$, ${}^{137}\text{Cs}$ and ${}^{212}\text{Pb}$, whereas the results for ${}^{22}\text{Na}$ show a slight deviation. As shown in Fig. 1, the spectrum of ${}^{22}\text{Na}$ has a significant continuum, which is impacted by the background. It should be noted that the coverage interval assessment with Fisher information matrix has limited performance when the number of counts is low. This is further confirmed with results of NaI simulations shown in Table 5, in which the number of counts of radionuclides is low. In particular, with a number of counts of only 500, the coverage interval evaluation of ${}^{134}\text{Cs}$ from Fisher information matrix has limited performance. The limited performances of using the Fisher information

matrix is due to: i), the larger estimation bias at low statistics which is shown in recent studies (Xu et al., 2020). ii), the underlying optimization problem of the spectral unmixing is a maximum likelihood estimation with non-negativity constraint which makes the Gaussian distribution no longer true at low statistics.

6. Discussion about the limitations of the proposed methods

In this paragraph, we discuss some potential limitations of the proposed approach.

Statistical limitation

The derivations of the detection threshold and the coverage interval are both based on assumptions that we discuss hereafter:

- *Computation of the decision threshold* : In the proposed approach, the detection threshold for a given radionuclide is defined from the statistical distribution of some test statistics T that depends upon an equivalent background. This equivalent background is function of the contribution of all the radionuclides except the tested one. Since the such an equivalent background depends upon the result of the unmixing process, it is implicitly a function of the *estimated* activity of all the radionuclides. However, it is important to highlight that this dependency is weak. Indeed, for each radionuclide to be tested, the equivalent background derives from the (weighted) sum of the contributions of all radionuclides except the tested one. These linear combinations are dominated by the radionuclides with the highest activities, which are generally accurately estimated thus providing a reliable equivalent background estimation.

Traditional peak-based methods also depend on the estimation of an equivalent background at the vicinity of the energy peaks of the tested radionuclides. In the proposed method, the equivalent background estimates bear more physical information as the method allows to account for the Compton continuum of each radionuclide in the full observed energy range.

- *Determination of the coverage interval* In the context of spectral unmixing, coverage intervals can be determined in several ways. One standard way would be to perform Monte-Carlo simulations, assuming some prior model about the mixture (*e.g.* the maximum likelihood estimate of the mixture weights). Similarly to the previous point dedicated to the detection threshold, the resulting coverage intervals will weakly depend on the estimates of the mixing weights. While Monte-Carlo simulations allow to determine potentially complex, non-symmetric, coverage intervals, they also require a significantly large computational cost to derive reliable coverage intervals, which might limit its applicability for fast analysis purposes. To that end, we rather proposed to opt for an approximation of the distribution of the mixing weights distribution based on the Fisher information matrix. The inverse of the Fisher information matrix provably provide the asymptotic covariance matrix of the mixing weights distribution about the maximum likelihood estimate. The quality of this approximation will strongly depend on the actual statistics of the data, at very low statistics, this approximation is unlikely to be valid. Similarly, for radionuclides with very low activities, the corresponding coverage intervals are likely to be non-symmetric.

Determining coverage intervals only make sense when for radionuclides that have activities larger than the detection thresholds. In the aforementioned limit cases, we made the assumption that the measured activities are high enough so that the coverage interval does not include 0. This is generally not the case for radionuclides whose activities are close to the detection threshold. In this case, the non-negativity constraint used in the spectral unmixing method prevents the estimated mixing weights to be negative. Consequently, the distribution of the estimated mixing weights can be described by a truncated Gaussian distribution defined as follows:

$$f(\theta) = p\delta(0) + \Pi_+(\theta) \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp - \frac{(\theta - a_i^{ML})^2}{2\sigma_i^2} \quad (29)$$

where $\delta(0)$ is the Dirac function centered about 0 and the probability weight at 0 is defined as:

$$p = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp -\frac{(\theta - a_i^{ML})^2}{2\sigma_i^2} d\theta$$

The function $\Pi_+(\theta)$ is a step function that takes the value 0 when $\theta < 0$ and 1 when $\theta > 0$. The coverage interval for a given probability α can then be defined as the smallest interval $[\sigma_1, \sigma_2]$ so that:

$$\int_{\sigma_1}^{\sigma_2} f(\theta)d\theta = \alpha \tag{30}$$

As an illustration, we perform a comparison of the above analytic model with Monte-Carlo simulations for ^{137}Cs with an activity taken at the detection threshold. Figure 14 shows the distribution obtained from 15000 Monte-Carlo simulations in blue and the analytic Fisher-based approximate distribution defined in 29, which show a very good agreement. This experiments highlights that the Fisher-based approximate distribution can be generalized to low-level cases, close to the detection threshold.

Modeling limitations

[AVO: A RELIRE. TROP LONG ???]

In standard gamma-ray spectrometry analysis the detection efficiency calibration is required for each radionuclide and each detection configuration (specific detector and counting geometry). Alike the determination of the spectral signatures is required for each radionuclide and each detection configuration in the spectral unmixing approach. Due to the unavailability of standard sources containing all the radionuclides of interest in single emitter standard source, the spectral signatures are calculated using Monte Carlo type particle transport codes. MCNP-CP simulates the interaction between the photons emitted by the radionuclide of the source (or the sample) and the germanium crystal, and the surrounding materials. In this paper the simulations were performed considering the detection system used for the aerosol filters measurements : the same

HPGe detector (61% Mirion-Canberra Broad Energy Germanium detector), the same counting geometry (10 mL) and the same sample matrix (polypropylene fibers). The sample height or position variability will be taken into account in the uncertainty budget described in the associated paper Xu et al. (2021) Jiabin et al. (2021). Another component is also added in the uncertainty budget to account for the simulation process : indeed the modeling of the detection configuration is adjusted to reproduce the full energy peak (FEP) efficiency which is experimentally determined by the measurement of a standard source whose activities are given with uncertainties. Neither the total detection efficiency nor the full spectrum shape can be compared to experimental data, except with single radionuclide source measurements. This lack of knowledge is certainly a limitation of the proposed approach in the case when the spectral signatures must be determined by simulation. Forthcoming study will be dedicated to the potential differences between simulated and experimental spectral signatures, using single gamma-ray emitter standard sources.

7. Conclusion

In this paper, we investigate how the characteristic limits, *i.e.* the decision thresholds and coverage intervals, can be determined for spectral unmixing algorithms. These two quantities, which are related to the statistical limits for decision making purposes in gamma ray spectra analysis, are first investigated for such full spectrum analysis methods.

Firstly, we propose to determine the decision threshold of each radionuclide based on different statistical tests of its equivalent background. This background spectrum, consisting of other spectral contributions provides a mean value of the distribution under null hypothesis, allowing quantifying the decision threshold based on a given false negative rate. Next, we propose to assess the limits of coverage intervals with the Fisher information matrix of estimated mixing weights, which approximates the standard deviation of the estimation distribution. This statistical uncertainty must be added to other uncertainty

components which are introduced and described in Xu et al. (2021) to obtain the final uncertainty.

The proposed approaches have been evaluated with Monte Carlo simulations of both HPGe and NaI detectors. They have been shown as efficient and reliable methods to both of these simulated data. As the simulated spectra are calculated for radionuclides in samples measured on the HPGe detector of the routine aerosols measurements in the laboratory, the methods proposed in this paper can be used to provide metrological results of experimental data, which is presented in Xu et al. (2021).

References

- André, R., Bobin, C., Bobin, J., Xu, J., de Vismes Ott, A., 2020. Metrological approach of γ -emitting radionuclides identification at low statistics: application of sparse spectral unmixing to scintillation detectors. *Metrologia* .
- Boshkova, T., 2018. Decision threshold and relative uncertainty of the measured net signal in radiation measurement. *Applied Radiation and Isotopes* 139, 146 – 150.
- Fisher, R.A., 1956. *Statistical methods and scientific inference*. Oliver and Boyd.
- ISO 11929, 2010. Determination of the characteristic limits (decision threshold, detection limit and limits of the confidence interval) for measurements of ionizing radiation - fundamentals and application.
- Korun, M., Vodenik, B., Zorko, B., 2014. Calculation of the decision thresholds in gamma-ray spectrometry. *Applied Radiation and Isotopes* 94, 221 – 229.
- Michel, R., 2016. Measuring, estimating, and deciding under uncertainty. *Applied Radiation and Isotopes* 109, 6–11.

- Mirion Technologies (Canberra), 2013. Genie 2000 customization tool manual v3.4.
- Paradis, H., Bobin, C., Bobin, J., Bouchard, J., Lourenço, V., Thiam, C., André, R., Ferreux, L., de Vismes Ott, A., Thévenin, M., 2020. Spectral unmixing applied to fast identification of γ -emitting radionuclides using NaI(Tl) detectors. *Applied Radiation and Isotopes* 158, 109068.
- Weise, K., Hübel, K., Michel, R., Rose, E., Schläger, M., Schrammel, D., Täschner, M., 2005. Determination of the detection limit and decision threshold for ionizing-radiation measurements: fundamentals and particular applications ISSN 1013-4506.
- Xu, J., Bobin, J., de Vismes Ott, A., Bobin, C., 2020. Sparse spectral unmixing for activity estimation in γ -ray spectrometry applied to environmental measurements. *Applied Radiation and Isotopes* 156, 108903.
- Xu, J., Bobin, J., De Vismes Ott, A., Bobin, C., Malfrait, P., 2021. Analysis of gamma-ray spectra with spectral unmixing -Part II: Calibrations for the quantitative analysis of HPGe measurements. URL: <https://hal.archives-ouvertes.fr/hal-03238923>. working paper or preprint.

	H_0 is true	H_1 is true
rejecting H_0	Type I error: the error of rejecting H_0 when it is true, the probability of committing a type I error is denoted by α , called false positive rate.	
accepting H_0	Type II error: the error of accepting H_0 when H_1 is true, the probability of committing a type II error is denoted by β , called false negative rate.	

Table 1: Two types of errors of hypotheses test.

${}^7\text{Be}$	${}^{22}\text{Na}$	${}^{40}\text{K}$	${}^{137}\text{Cs}$	${}^{210}\text{Pb}$	${}^{208}\text{Tl}$	${}^{212}\text{Bi}$	${}^{212}\text{Pb}$	${}^{214}\text{Bi}$	${}^{214}\text{Pb}$
391	0.078	1.24	0.015	27	0.098	0.425	0.239	0.054	0.056

Table 2: Activity (Bq) of radionuclides of HPGe simulations.

Radionuclide	${}^{60}\text{Co}$	${}^{134}\text{Cs}$	${}^{137}\text{Cs}$	${}^{152}\text{Eu}$	background
Number of counts	1500	500	1500	2500	4000

Table 3: Number of counts of NaI simulations.

	${}^7\text{Be}$	${}^{22}\text{Na}$	${}^{137}\text{Cs}$	${}^{212}\text{Pb}$	expected value
percentage within $a_0 \pm \sigma_f$	68.0	67.3	68.87	68.25	68.27
percentage within $a_0 \pm 2\sigma_f$	96.49	94.28	95.85	95.02	95.45
percentage within $a_0 \pm 3\sigma_f$	99.9	99.8	99.76	99.9	99.73

Table 4: Standard deviation from Fisher information matrix comparing to Monte Carlo simulations (HPGe measurements).

	^{60}Co	^{134}Cs	^{137}Cs	^{152}Eu	expected value
percentage within $a_0 \pm \sigma_f$	68.38	65.81	68.7	65.58	68.27
percentage within $a_0 \pm 2\sigma_f$	95.34	94.78	94.62	95.23	95.45
percentage within $a_0 \pm 3\sigma_f$	99.73	99.54	99.95	99.8	99.73

Table 5: Standard deviation from Fisher information matrix comparing to Monte Carlo simulations (NaI measurements).

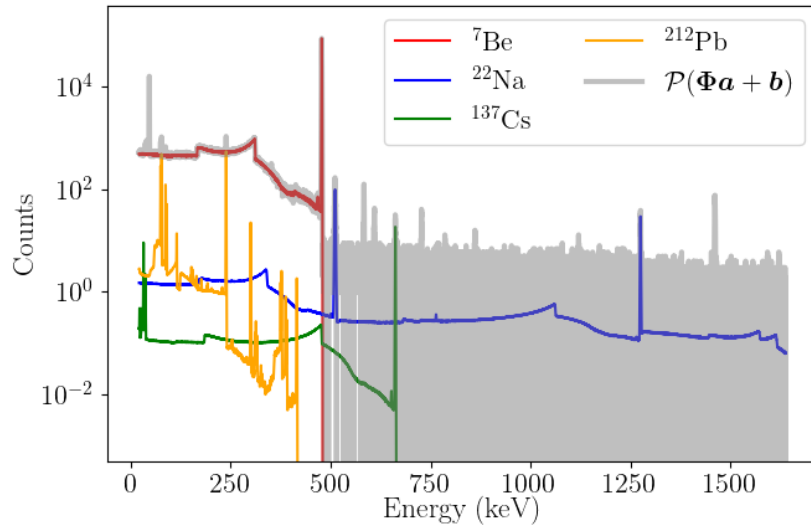


Fig. 1: Spectral unmixing model used to illustrate the evaluation of the characteristic limits. The simulated measurements (gray) are composed of 10 radionuclides: ^7Be , ^{22}Na , ^{40}K , ^{137}Cs , ^{210}Pb , ^{208}Tl , ^{212}Bi , ^{212}Pb , ^{214}Bi , ^{214}Pb , while individual spectra of ^7Be , ^{22}Na , ^{137}Cs , ^{212}Pb will be evaluated.

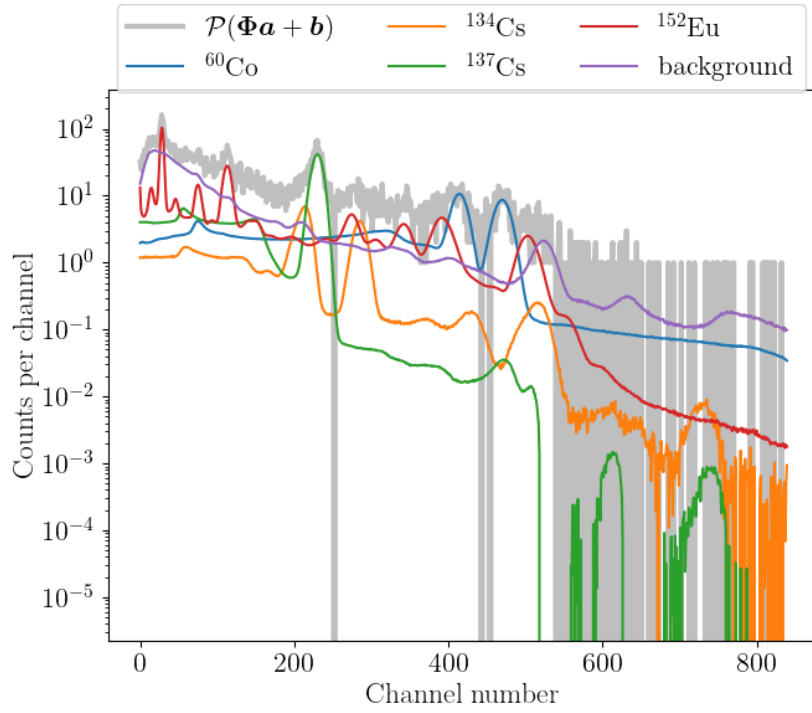


Fig. 2: Spectral unmixing model of simulations of NaI measurements.

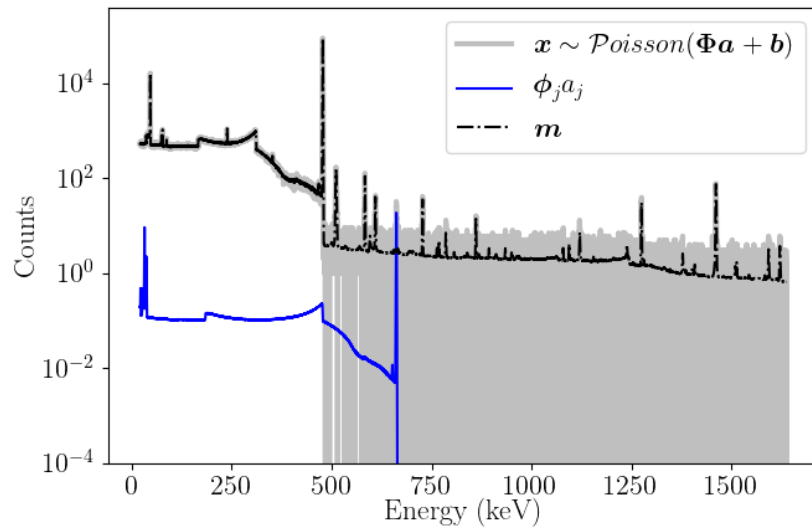


Fig. 3: Illustration of equivalent background spectrum of the ^{137}Cs .

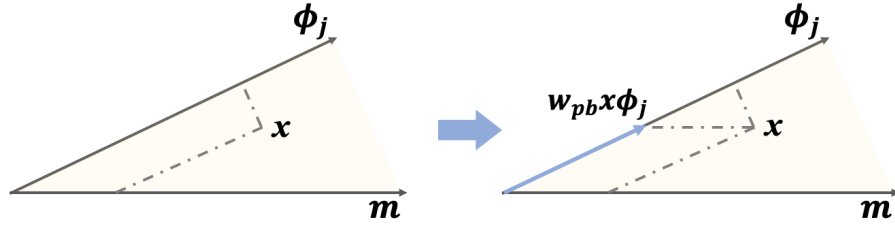


Fig. 4: [JB: Figure à refaire, la projection sur m ne devrait pas être // à gauche] ist of parallel projection with respect to the equivalent background. The vector ϕ_j is the spectral signature of the radionuclide. The vectors \mathbf{x} , \mathbf{m} are the measured spectrum and the equivalent background spectrum. w_{pb} is the weight vector.

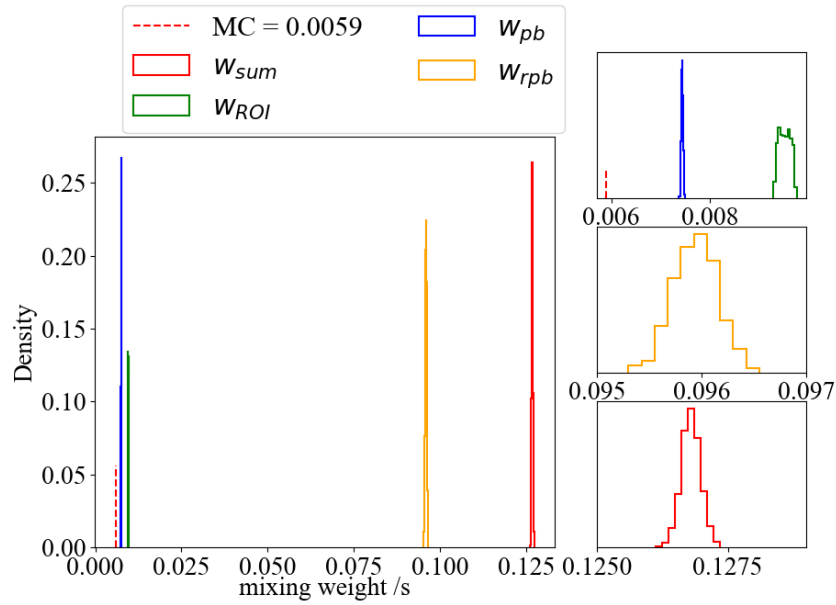


Fig. 5: Decision threshold assessment for HPGe measurements: ${}^7\text{Be}$. In Fig. 5, 6, 7 and 8: the right plot shows the distribution of calculated DT based on the statistical test of equivalent background weighted by w_{sum} (red), w_{ROI} (green), w_{pb} (blue) and w_{rpb} (orange). The accurate DT level calculated from Monte Carlo (*i.e.*, $1 - \alpha$ percentile of the distribution under the null hypothesis) is displayed with a dotted line (red). The left panels show the zoom of the distributions in the right plot.

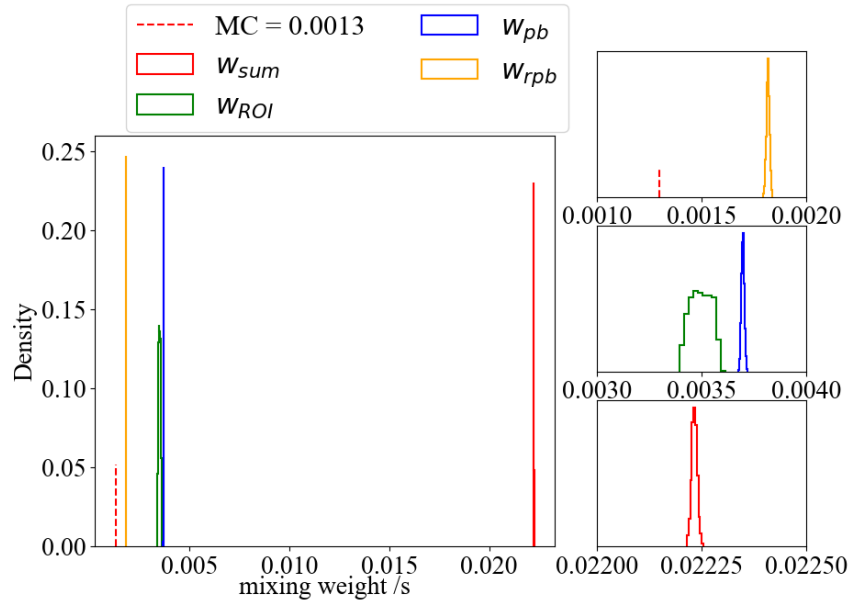


Fig. 6: Decision threshold assessment for HPGe measurements: ^{22}Na .

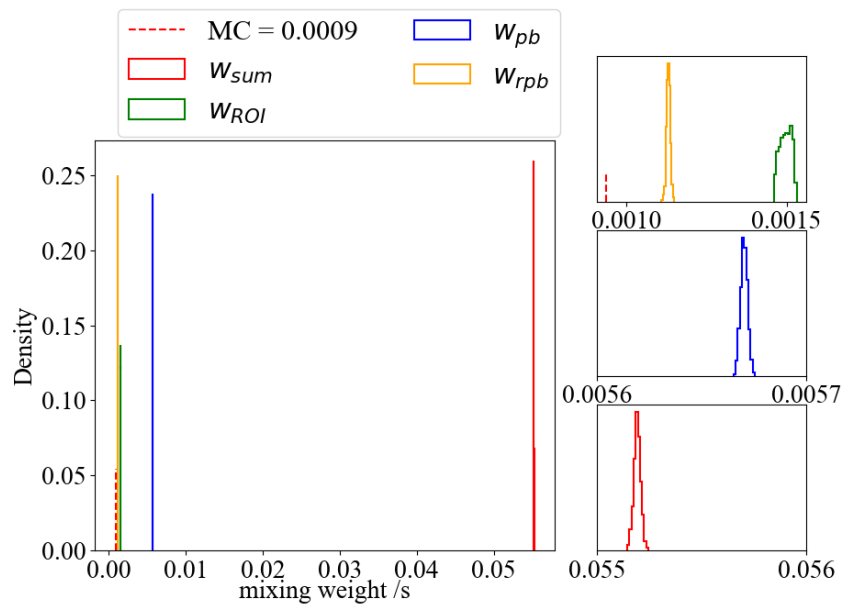


Fig. 7: Decision threshold assessment for HPGe measurements: ^{137}Cs .

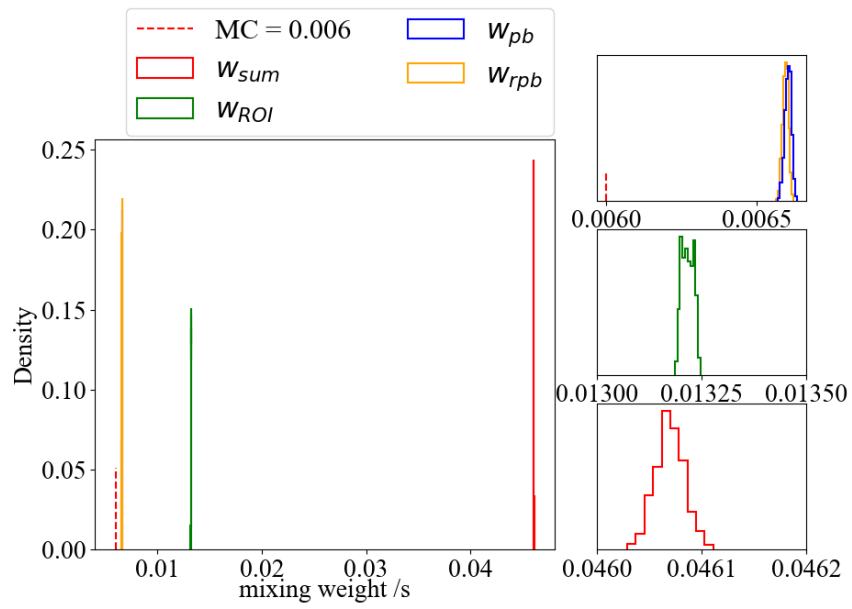


Fig. 8: Decision threshold assessment for HPGe measurements: ^{212}Pb .

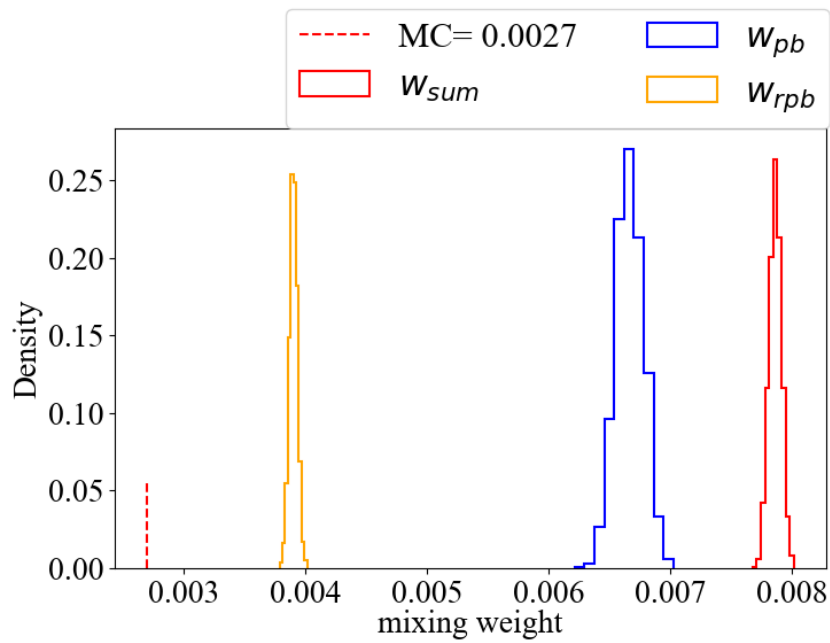


Fig. 9: Decision threshold assessment for NaI measurements: ^{60}Co . Fig. 9, 10, 11 and 12: the plot shows the distribution of calculated DT based on the statistical test of equivalent background weighted by w_{sum} (red), w_{pb} (blue) and w_{rpb} (orange). The accurate DT level calculated from Monte Carlo (*i.e.*, $1-\alpha$ percentile of the distribution under the null hypothesis) is displayed with a dotted line (red).

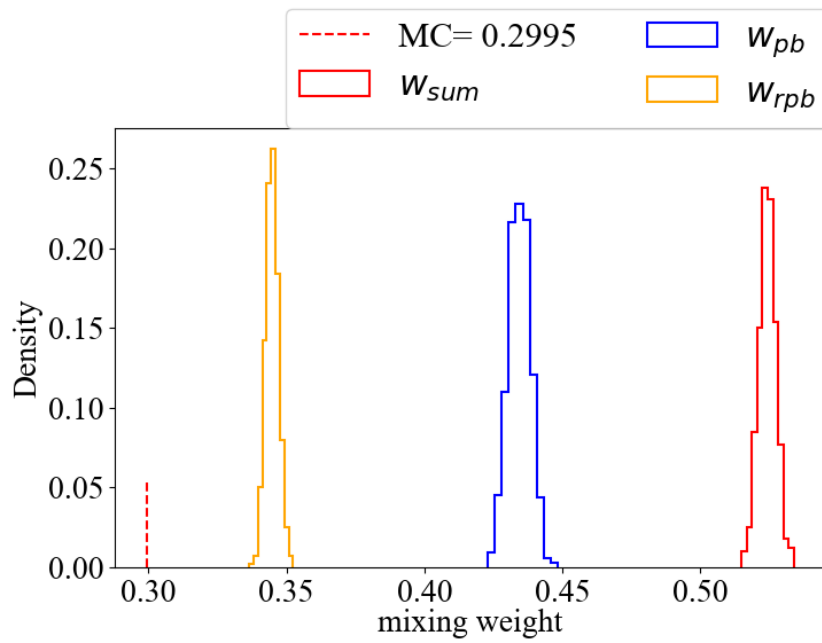


Fig. 10: Decision threshold assessment for NaI measurements: ^{134}Cs .

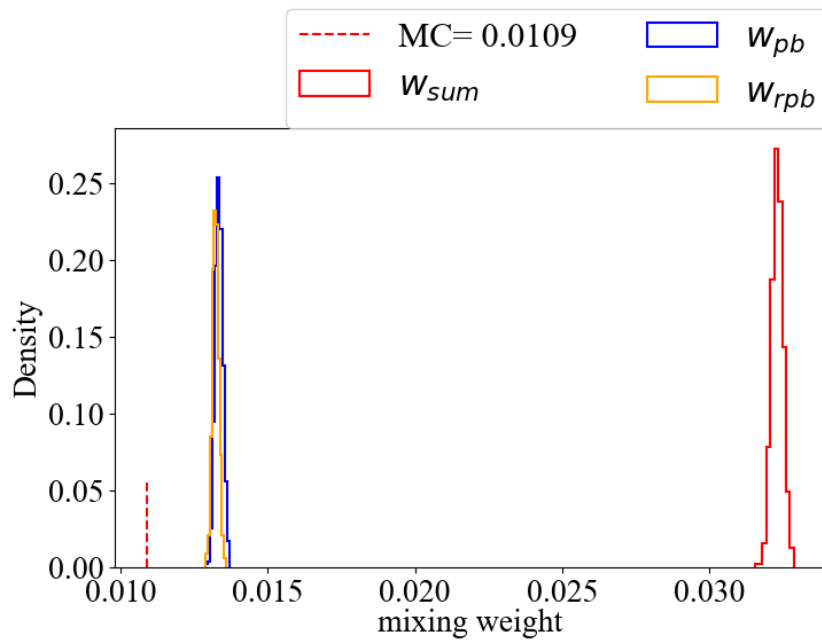


Fig. 11: Decision threshold assessment for NaI measurements: ^{137}Cs .

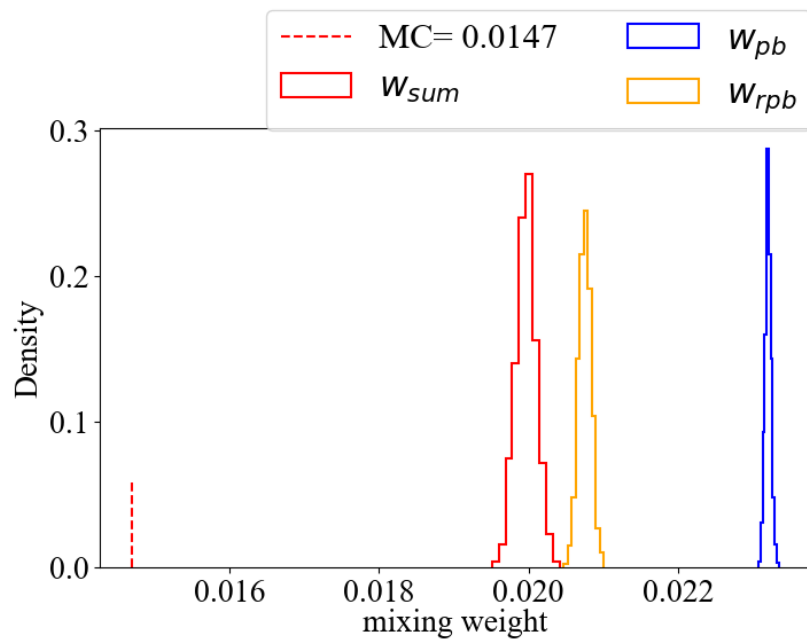
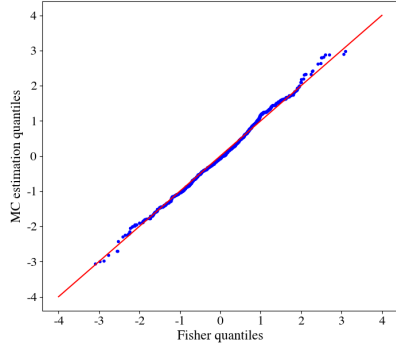
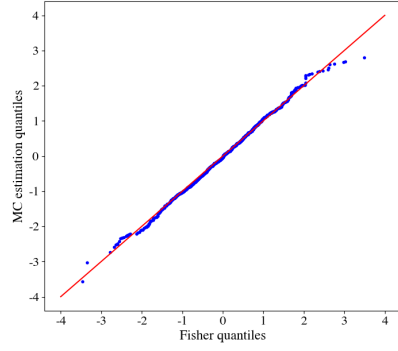


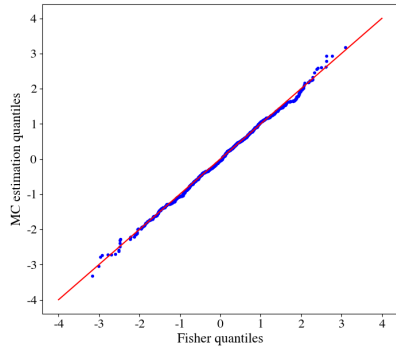
Fig. 12: Decision threshold assessment for NaI measurements: ^{152}Eu .



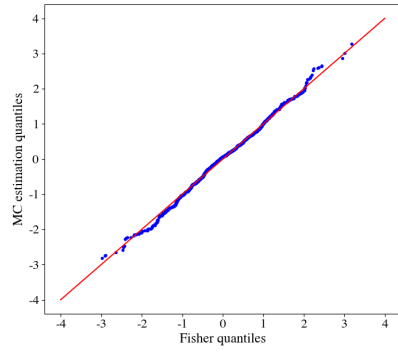
(a) ${}^7\text{Be}$



(b) ${}^{22}\text{Na}$



(c) ${}^{137}\text{Cs}$



(d) ${}^{212}\text{Pb}$

Fig. 13: Evaluation for the HPGe simulation: the Q-Q plots for Fisher quantiles ($\mathcal{N}(a_0, \bar{\sigma}_f^2)$, where a_0 is the expected mixing weight of the radionuclide and $\bar{\sigma}_f$ is the mean value of the standard uncertainty calculated from the Fisher information matrix of estimated mixing weights of 1000 Monte Carlo simulations.) versus estimation distribution quantiles for radionuclides (blue). The theoretical Q-Q plots are shown in red.

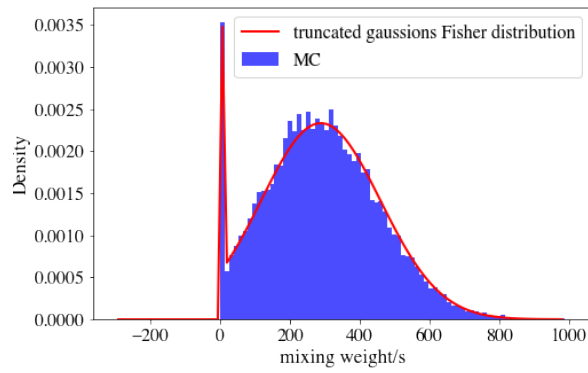
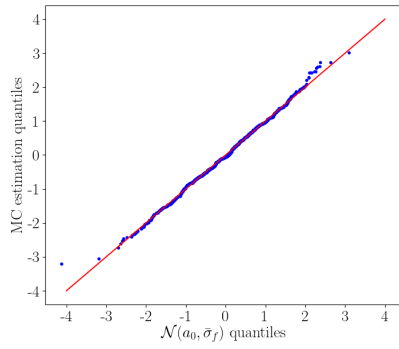
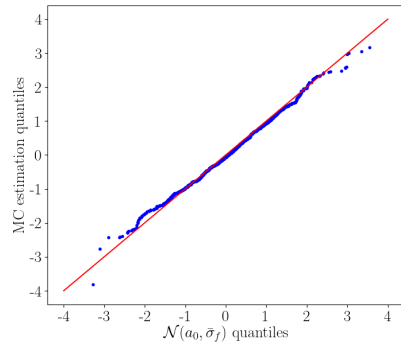


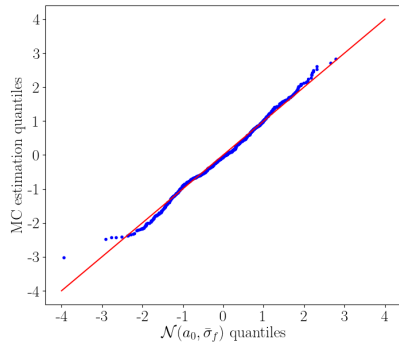
Fig. 14: Evaluation of coverage intervals assessment with Fisher information matrix: HPGe measurements with ^{137}Cs at decision threshold level. The distribution of estimated mixing weight of ^{137}Cs of 15000 Monte Carlo simulations (blue) is compared to the probability density function of the truncated normal distribution $\mathcal{N}(a_0, \sigma_f)$.



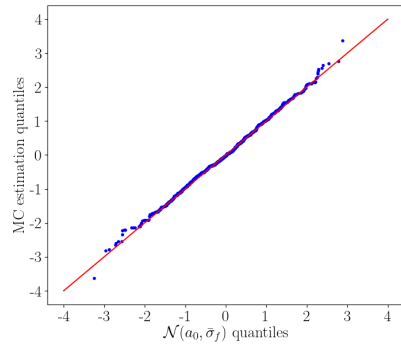
(a) ^{60}Co



(b) ^{134}Cs



(c) ^{137}Cs



(d) ^{152}Eu

Fig. 15: Evaluation for the NaI simulation: the Q-Q plots for Fisher quantiles versus estimation distribution quantiles for radionuclides.(blue). The theoretical Q-Q plots are shown in red.