Automatic identification and quantification of γ-emitting radionuclides with spectral variability using a hybrid Machine Learning unmixing method

Dinh Triem Phan[a,*], Jérôme Bobin[b], Cheick Thiam[a], Christophe Bobin[a]

[a] *Université Paris-Saclay, CEA, List, Laboratoire national Henri Becquerel (LNE-LNHB), Palaiseau, 91120, France*
[b] *IRFU, CEA, Universite Paris-Saclay, Gif-sur-Yvette, 91191, France*

Abstract

Automatic identification and quantification of γ-emitting radionuclides, considering the spectral deformation due to γ-interactions in the radioactive source environment, is a demanding challenge in the field of nuclear physics. In this context, this paper presents a hybrid unmixing approach combining a pre-trained machine learning model (autoencoder) to capture spectral deformations and a model selection technique based on statistical testing to identify the radionuclides present in a radioactive source and quantify their mixing weights. The identification process of radionuclides is based on a sequential selection algorithm using a likelihood ratio test depending on the expected false positives. Basically, this method aims to minimize the number of radionuclides in an initial radionuclide library containing characteristic γ-spectra of each γ-emitter to be tested. The robustness of decision-making with this approach and the quantification performance are investigated with Monte Carlo simulations involving up to 12 radionuclides to be tested, according to different mixture scenarios with increasing complexity and various statistics. The false positive rates obtained with the hybrid unmixing approach are close to the expected values. In general, the quantification results are similar to the case when the radionuclides present in the source and spectral signatures are known. This highlights the effectiveness of this novel hybrid unmixing approach for the automatic identification and quantification of gamma-ray emitting radionuclides with spectral variability.

Keywords: Gamma-ray spectrometry, Automatic identification and quantification, Spectral variability, Hybrid algorithm, Machine Learning, Model selection

*Corresponding author
Email address: dinh-triem.phan@cea.fr (Dinh Triem Phan)

1. Introduction

Gamma-ray spectrometry stands as a traditional technique for identifying and quantifying γ-emitting radionuclides in many applications in nuclear physics. The traditional γ-spectrum analysis employed by experts is based on the peak-fitting technique applied to full-absorption peaks [1] assuming that the measurement noise follows a Gaussian distribution. Gamma-spectrometry faces several challenges for specific applications, such as rapid identification to prevent illegal trafficking of nuclear materials [2], decommissioning of nuclear facilities or in situ environmental analysis following a radiological or nuclear accident. For these applications, it is crucial to have a method that can be operated by non-expert users while enabling robust decision-making (*e.g.* minimizing false alarms to limit expert intervention) under short measurement conditions of a few seconds or minutes (*i.e.* low statistics) with the possibility of estimating radionuclide activity. Since traditional methods are not well-suited to address these challenges, Machine Learning (ML) algorithms have been recently investigated [3–7]. However, ML approaches do not allow to explicitly account for physical knowledge about the measurement process (e.g. mixture model, measurement statistics, etc.), which may hamper their precision. Furthermore, the size of the training set grows factorially with the complexity of the mixture, especially with respect to the number of radionuclides to be tested. Besides these ML developments, statistical approaches to γ-spectrometry provide good results in terms of accuracy and uncertainty assessment. In this context, full-spectrum unmixing analysis based on maximum likelihood of Poisson statistics and a model selection technique yields effective performances for robust identification and accurate quantification of γ-emitting radionuclides at low statistics for HPGe and NaI(Tl) detectors [8, 9]. For instance, experimental results on real aerosol filter samples show that this method gives faster results for the detection of ^{137}Cs with very low activity (mBq), around a day and a half after sampling, whereas traditional methods for peak area determination require around eight days [10, 11].

In practical applications, the measurements carried out in complex environments can lead to variability of the shape of γ-spectra due to physical phenomena such as Compton scattering, attenuation, or fluorescence. While some studies have explored this problem using analytical formulas, simulations [12, 13] or ML [4, 14], these approaches are not well adapted for the context of automatic identification or precisely account for the spectral variability as well as Poisson statistics. The work in [15] introduced a hybrid ML unmixing algorithm that can accurately take into account these issues for the quantitative analysis of γ-spectra. This algorithm combines a pre-trained ML model based on an autoencoder architecture that captures the spectral deformation. This algorithm, coined SEMSUN, performs a maximization of a regularized likelihood estimation (MLE) of both the mixing weights (counting) of the radionuclides and their spectral signatures. In this work, the radionuclides present in the radioactive source are supposed to be known.

In the case where no spectral variability is present, previous work [8, 9] have

proposed a dedicated full-spectrum analysis method, the Poisson-Orthogonal Matching Pursuit (P-OMP) algorithm, which can robustly identify the γ-emitting radionuclides from a library of radionuclides. This algorithm is based on a sequential model selection procedure using a likelihood ratio test (LRT) [16] to identify the radionuclides that contribute to experimental γ-spectra.

The present work aims to combine the P-OMP and SEMSUN algorithms to tackle radionuclide identification and quantification when spectral variability is present. To that end, this paper investigates the model selection P-OMP algorithm in combination with the hybrid ML algorithm to identify and quantify γ-emitting radionuclides. Identification performance is evaluated in terms of false positives and false negatives to meet the demands of certain applications such as illegal trafficking or environmental analysis.

The article is organized as follows:

- Section 2 describes the P-OMP model selection algorithm used to identify the radionuclides contributing to an observed γ-spectrum without spectral variability.
- The combination of the hybrid ML and the P-OMP algorithms for radionuclide identification with spectral variability is investigated in Section 3.
- Numerical assessment for radionuclide identification and quantification is presented in Section 4 based on various mixing scenarios of radionuclides and different statistics of γ-spectrum. A library contains 12 radionuclides covering a large range of energies between 20 keV and 2 MeV. Most of these radionuclides are classical γ-emitters for handheld instruments or spectroscopy-based portal monitors (*e.g.*, ANSI N42.38-2015 [17]). Specific tests are also carried out when the selection model algorithm can lead to confusion between identified radionuclides having correlated γ-spectra.

2. Identification of radionuclides with fixed spectral signatures (without variability)

2.1. Radionuclide identification as a model selection problem

An observed γ-spectrum $y = [y_1, ..., y_M]$ is a vector of M channels, with each channel representing energy intervals registered by a detector sensitive to γ-photons. For each channel m, the counting y_m is the number of events that the deposited energy fell within that interval. The observed spectrum y follows the Poisson distribution of Xa: $y \sim \mathcal{P}(Xa)$. X is a matrix composed of the normalized spectral signature of all radionuclides (including the natural background (Bkg)), and a is the vector containing the counting or mixing weights of each radionuclide. Spectral unmixing for γ-spectrum analysis is an inverse problem applied to estimate the counting a from an observed spectrum y and the spectral signatures X. The cost function to be minimized is the negative logarithm of the likelihood (equivalent to the divergence), which has the following

formula:

$$D(y||Xa) = \sum_{m=1}^{M} ((Xa)_m - y_m + y_m log(y_m) - y_m log((Xa)_m)) \quad (1)$$

Practically, X is a matrix containing all the spectral signatures of radionuclides to be tested during the unmixing process. When the actual number of radionuclides present in a given sample is not known in advance, the number of radionuclides to be tested can be large compared to the actual value. In this context, minimizing the cost function of Eq 1 will likely lead to noisy solutions. Indeed, additional radionuclides, which are actually not present in the sample, will tend to overfit the noise of the observed spectrum. From an identification perspective, this will yield false detections.

In order to overcome these issues, extra regularisation is required. Spectral unmixing with automatic identification can then be considered from a model selection perspective. In the present case, this can be tackled by selecting a regression model (*i.e.* the set of detected radionuclides) with minimal dimensionality. In the statistics literature, various approaches for model selection with dimension penalization have been proposed, with a strong focus on using information criteria (Akaike Information criterion and Bayes Information criterion, to name only two). However, they generally lack an automatic model selection procedure to precisely select a model for a given detection performance level.

In [8, 9], another approach consists in enforcing the sparsity of the counting vector a. Penalization of the number of identified radionuclides is equivalently reformulated as penalizing the non-zero entries of the counting vector a. To that end, the optimization problem can be recast as follows:

$$\hat{a} = \underset{a}{\operatorname{argmin}}\, \chi(. \geq 0)(a) + D(y||Xa) \text{ subject to } ||a||_0 \leq K \quad (2)$$

where $\chi_{(.\geq 0)}(a)$ represents the positivity constraint on the vector a, which equals to 0 when a has no negative entry and $+\infty$ otherwise. $||a||_0$ represents the number of non-zero elements in a, which enforces the number of identified radionuclides to be less than K.

2.2. P-OMP: a model selection algorithm for γ-ray spectrometry

The problem in Eq 2 is combinatorial in nature. In the reference [18], the authors introduced a greedy minimization algorithm to tackle this problem with a predefined number of identified radionuclides. However, the number of radionuclides present in the source is generally unknown. More recently, the works in [8, 9] proposed implicitly imposing sparsity thanks to a greedy model selection procedure, where models with increasing dimensionality are tested, and a statistics-based criterion to stop the selection procedure.

The Poisson Orthogonal Matching Pursuit algorithm (P-OMP - [8, 9]) performs sequentially by tackling unmixing problems with a fixed number of identified radionuclides:

$$\hat{a} = \underset{a}{\operatorname{argmin}}\, \chi(. \geq 0)(a) + D(y||Xa) \text{ subject to } ||a||_0 = K \quad (3)$$

The P-OMP algorithm starts from a single radionuclide to be added to the natural background, which is always present in the measurement. In that case, this problem is equivalent to searching for the radionuclide that provides the highest likelihood using the non-negative Poisson unmixing algorithm (NNPU) [8]. At each iteration, a single radionuclide is added to the library X; the model selection procedure then consists of picking the radionuclide that leads to the largest increase in the likelihood.

The stopping criterion is based on a statistical hypothesis testing procedure using the LRT [16] between the current model and the new model with an additional radionuclide. If the current model is correct and the new model includes an extra, non-present radionuclide, the negative log-likelihood difference between these two models (called deviance [19]) asymptotically follows a Chi-square distribution with a number of degrees of freedom that is equal to the number of extra parameters in the new model with respect to the current model (one in this setting). The P-OMP algorithm then stops when the deviance is below a threshold that corresponds to a given p-value α of the asymptotic Chi-square distribution.

From a statistical perspective, several radionuclides are to be tested at each step, which would require a cumbersome multiple-hypothesis testing procedure. Rather, the P-OMP performs parallel single-hypothesis tests, which are known to be biased. To that end, the Bonferroni correction [20] is used to correct the test by dividing the expected false positive rate α by the number of radionuclides to be tested in the library.

In addition, it has been proposed in [8] to correct the calculation of the decision threshold. According to the maximum likelihood property, the estimated counting of each radionuclide asymptotically follows a Gaussian distribution with a mean value equal to the expected counting. For the radionuclide not present in the radioactive source, the estimated counting asymptotically follows a centered normal distribution. Due to the positivity constraint on the counting, all the negative values become equal to 0. For these values, no false positive is detected by the algorithm since the deviance is close to 0 and always lower than the threshold. Hence, the false positive rate in this case is two times lower than the expected value. To ensure the same false positive rate, the expected value must be multiplied by two.

The pseudocode for this algorithm is described in [8, 9]. The schema of P-OMP is illustrated in Fig. 1.

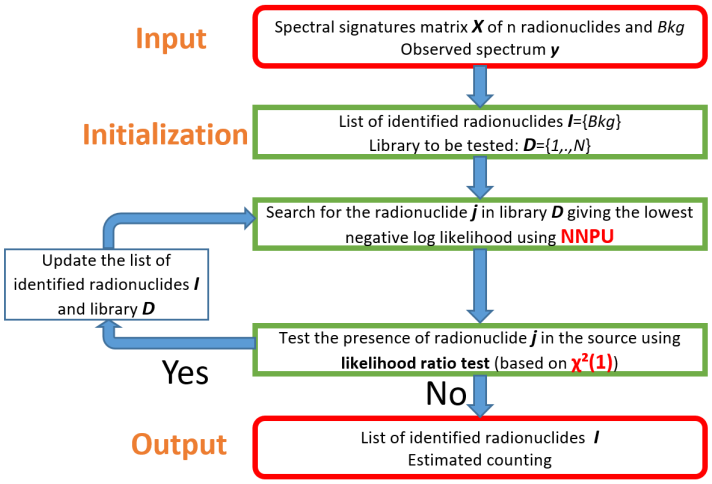

Fig. 1. Schema of P-OMP.

3. Identification of radionuclides with spectral variability

As emphasized in [15], dealing with spectral variability requires adapting the spectral unmixing algorithm to jointly estimate the counting of the radionuclides and their spectral signatures. This section presents how the algorithm proposed in [15], Semi-blind Spectral unmixing based on manifold learning (SEMSUN), can be generalized further to identify the unknown radionuclides in the presence of spectral variability.

3.1. Semi-blind Spectral unmixing based on manifold learning (SEMSUN)

In [15], a hybrid unmixing algorithm has been introduced that combines an MLE-based statistical approach to model measurement statistics and ML-based models to capture spectral variability. More precisely, a specific ML model, called IAE (Interpolating AutoEncoder [21]), is learned to jointly model the spectral variability of all the radionuclides, thus providing a generative model of the latter. Hence, the spectral signatures can be described by a pre-trained nonlinear function, parameterized with a low-dimensional latent variable λ. This IAE-based model can be used as a surrogate model of the spectral signatures in the unmixing process to estimate jointly the spectral signatures and the radionuclide counting.

In the proposed framework, the spectral signatures X can be described as a known function of the latent variable λ, denoted $f(\lambda)$. The problem is to estimate the spectral signature X (or the latent variable λ) and the counting vector a that minimizes the cost function:

$$\hat{\lambda}, \hat{a} = \underset{\lambda \in S_d^+, a}{\operatorname{argmin}} \chi(. \geq 0)(a) + D(y||X_{Bkg}a_1 + f(\lambda)a_{i>1}) \quad ; \quad \hat{X} = f(\hat{\lambda}) \qquad (4)$$

where $\chi_{(.\geq 0)}(a)$ represents the positivity constraint on the vector a and S_d^+ is the non-negative and simplex space. The function f stands for the decoder of encoded anchor points of the IAE (see [15] for more details about the IAE).

The above problem is not convex but is approximately multi-convex: it is convex with respect to the counting a when the latent variable λ is fixed and approximately convex as a function of λ once a is fixed. This type of optimization problem can be tackled with iterative and sequential algorithms such as the BCD algorithm (Block coordinate descent [22]), with two distinct blocks: one for a and one for X. The gist of the BCD consists of minimizing over each block, assuming the other one is fixed, in a sequential and iterative process:

- the counting of the radionuclides a^{k+1} is updated by minimizing the optimization problem 4 assuming X^k is fixed:

$$a^{k+1} = \underset{a}{\operatorname{argmin}}\, \chi(. \geq 0)(a) + D(y||X_{Bkg}a_1 + X^k a_{i>1}), \tag{5}$$

  which amounts to a standard spectral unmixing problem, solved using the NNPU algorithm [8].

- the spectral signatures X^{k+1} (or equivalently the latent variable of the IAE model λ^{k+1}) are updated for fixed counting a^{k+1}:

$$\lambda^{k+1} = \underset{\lambda \in S_d^+}{\operatorname{argmin}}\, D(y||X_{Bkg}a_1^{k+1} + f(\lambda)a_{i>1}^{k+1}) \quad ; \quad X^{k+1} = f(\lambda^{k+1}), \tag{6}$$

  which can be tackled using the numerical solver such as L-BFGS [23].

3.2. MoSeVa: model selection with spectral variability based on manifold learning

In [15], the set of radionuclides present in the measurements was assumed to be known and fixed, while both the spectral signatures and counting of the radionuclides were to be estimated. In the present context, the actual number of radionuclides to be present in the measurement is unknown and has to be selected from a potentially large library. Additionally, it is assumed that these radionuclides have spectral variability modeled with a pre-trained IAE-based surrogate model.

If the number of identified radionuclides K were known, the unmixing problem could be recast as follows:

$$\hat{\lambda}, \hat{a} = \underset{\lambda \in S_d^+, a}{\operatorname{argmin}}\, \chi(. \geq 0)(a) + D(y||X_{Bkg}a_1 + f(\lambda)a_{i>1}) \text{ subject to } ||a||_0 \leq K \tag{7}$$

where $f(\lambda)$ stands for the full library of radionuclide spectral signatures generated by IAE and $a_{i>1}$ is the corresponding counting vector. In the above problem, the number of identified radionuclides is imposed by looking at the

K radionuclides that yield the lowest negative logarithm likelihood or, equivalently, the highest likelihood, as testified by the pseudo-norm $||a||_0$ that counts the number of non-zero entries in a.

As this problem is combinatorial in nature, greedy sequential algorithms such as OMP or P-OMP have been introduced in signal processing to provide approximate but tractable solutions. Tackling model selection with spectral variability can be achieved by generalizing the P-OMP algorithm. The major difference lies in the use of the hybrid SEMSUN algorithm to tackle the spectral unmixing step for a fixed set of identified radionuclides.

A major discrepancy with the standard P-OMP algorithm lies in the model selection procedure. Indeed, when the spectral signatures are known, the deviance test used in the P-OMP to stop the selection procedure is theoretically guaranteed. However, when spectral variability is present, the distribution of the statistical deviance is more difficult to derive, which makes the theoretical guarantees unlikely to hold. This point will be illustrated in the next section.

The schema of MoSeVa is illustrated in Fig. 2. The pseudocode of this algorithm is presented in Appendix A. The data and codes are available on: `https://github.com/triem1998/MoSeVa`.

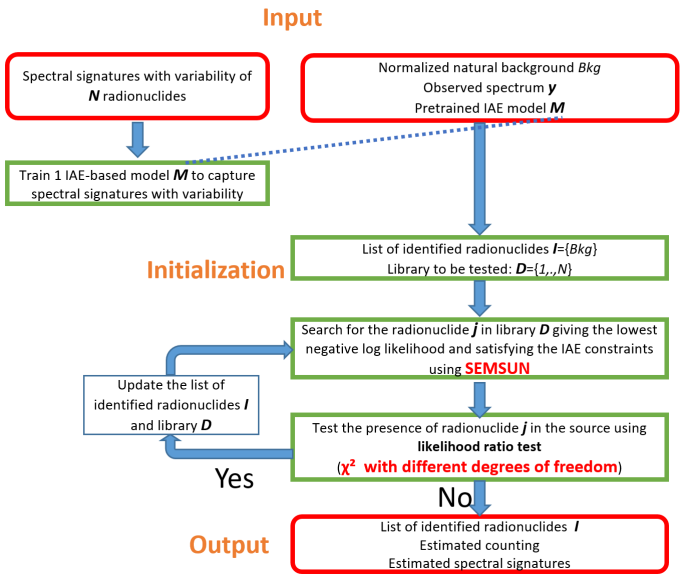

Fig. 2. Schema of MoSeVa.

4. Numerical evaluation of spectral unmixing

4.1. Data and evaluation metrics

The dataset for radionuclide identification and quantification is identical to that in the study [15]. Simulations of the spectral deformations were obtained

using Geant4 [24]-a Monte Carlo toolkit for simulating the transport of particles through matter. A point source is assumed to be located in a steel sphere and positioned 20 cm from the 3"×3" NaI(Tl) detector. For each radionuclide, different spectral signatures were generated with Geant4 simulations by varying the thickness of the steel sphere from 0.001 mm to 30 mm. In this case, the spectral signature of each radionuclide is mostly deformed by the attenuation and Compton scattering effects. The configuration of γ-ray spectrometry simulation on Geant4 is shown in Fig. 3.

The library was extended from 4 to 12 radionuclides covering a large range of energies between 20 keV and 2 MeV: ^{57}Co, ^{60}Co, ^{67}Ga, ^{88}Y, ^{99m}Tc, ^{123}I, ^{131}I, ^{133}Ba, ^{137}Cs, ^{152}Eu, ^{207}Bi and ^{241}Am. For each radionuclide, 96 spectral signatures were simulated by varying the steel thicknesses. In each Geant4 simulated spectrum, the emitted γ-photons follow the specific decay scheme for each radionuclide, and the total number of emitted γ-rays is around 250 million. The energy resolution of the NaI(Tl) detector was considered (6.5 keV at 662 keV). Each spectral signature consists of 1024 channels, with a binning of 2 keV per channel and a low-energy cut-off of 20 keV. The architecture and hyperparameters of IAE were the same as in the work [15]. One IAE model was trained to model the spectral signatures of 12 radionuclides with the same number of anchor points (the λ dimension remains one).

Radionuclide identification was investigated in the following sections with different radionuclide mixing scenarios. The false prediction rate, which includes the false negative rate (FNR) and the false positive rate (FPR), was used as a metric to assess the outcome of the identification process. False positives (FP) represent cases where at least one radionuclide is identified as present in the radioactive source when it is not. False negatives (FN) represent cases where at least one present radionuclide is undetected. For quantification results, the distributions of estimated counting a and spectral signatures X for all radionuclides were also studied.

Concerning the identification result, the new MoSeVa approach was compared to the P-OMP algorithm when spectral signatures are fixed and known. For quantification, this algorithm was compared to P-OMP to assess the impact of the joint estimation of spectral signatures and counting. This algorithm was also compared to SEMSUN (present radionuclides are assumed to be known) to assess the impact of the model selection. Oracle refers to the best possible result when the present radionuclides and their spectral signatures are known.

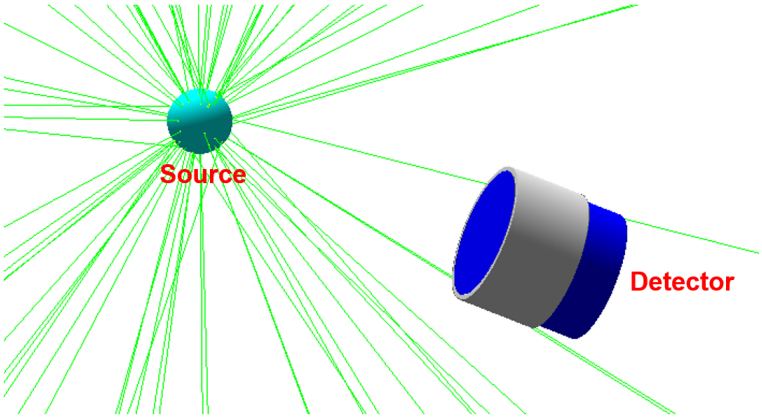

Fig. 3. Configuration of γ-ray spectrometry simulation on Geant4. A point source is located in a steel sphere and positioned 20 cm from the 3"×3" NaI(Tl) detector.

4.2. Investigation of the statistical deviance test

As indicated in 3.2, the deviance distribution is not determined theoretically. To further illustrate this point, one has to consider two distinct cases: the measured spectrum is only due to the natural background (*i.e.* no radionuclide is present), and the source contains at least one radionuclide.

4.2.1. No radionuclide is present

The first step of the MoSeVa algorithm boils down to testing the addition of a single radionuclide to a Bkg-only model. The MoSeVa algorithm then selects the radionuclide that provides the lowest cost function using the SEMSUN algorithm and checks if the new model passes the statistical deviance test. In contrast to the standard P-OMP, not only the counting but also the spectral signature of this radionuclide are estimated. With the IAE-based model, the spectral signatures of the radionuclides are characterized by a single latent variable λ with dimension d (number of anchor points, see [15, 21] for more details). The number of degrees of freedom of the latent variable is $d-1$ due to simplex constraint, which means that $d+1$ parameters must be estimated: the Bkg counting, the counting of this radionuclide and the latent variable λ. The hypothesis testing amounts to comparing the Bkg-only model with one that composes Bkg with the most likely radionuclide. The expected degrees of freedom in that case should be equal to d. However, the counting a and the latent variable λ are correlated, making the actual number of degrees of freedom lower than d and unknown. Theoretically, it can be demonstrated that using $\chi^2(1)$, the resulting FPR is higher than expected (see Appendix B). Applying $\chi^2(d)$ leads to a lower-than-expected value that is not theoretically proven. In the present context, only two anchor points were used so d equals 2.

The first numerical experiments were carried out using the tested radionuclide library containing only Bkg and a radionuclide i. Average Bkg counting

was evaluated in three cases: 50000, 5000 and 1250. For each average counting, 1000 Monte Carlo simulated γ-spectra were generated, each being a different simulated spectrum of Bkg.

Fig. 4 shows the FPR for 12 radionuclides using $\chi^2(1)$ and $\chi^2(2)$ when the average Bkg counting is equal to 1250. Generally, the FPR obtained when applying $\chi^2(1)$ is higher than expected, and when applying $\chi^2(2)$ is lower than expected. Note that the FPR in the cases where ^{241}Am and ^{123}I are tested is slightly lower than expected due to statistical noise when the number of Monte Carlo simulations is not sufficiently large. The radionuclide with more complex spectral variability (*e.g.*, ^{152}Eu, ^{133}Ba and ^{67}Ga) has larger FPR than radionuclides with limited variability (^{60}Co, ^{241}Am). Table 1 shows the result for different average Bkg counting with the expected FPR 5%. The same conclusion is reached in this case.

Subsequently, the same experiment was carried out with the library containing 12 radionuclides. In this case, it is a multiple-hypothesis test, so the selection procedure is performed, and the Bonferroni correction [20] is needed. Table 2 and Table 3 show the result when applying $\chi^2(1)$ and $\chi^2(2)$. Clearly, the FPR when using $\chi^2(1)$ is higher than expected, and when using $\chi^2(2)$ is lower than expected. It should be noted that when using $\chi^2(2)$, the decision threshold is larger, so the FNR obtained is also greater.

The choice between using $\chi^2(1)$ or $\chi^2(2)$ depends on the specific requirements of the application. When the objective is to maintain the FPR close to the expected value, and a slightly higher rate is tolerated, $\chi^2(1)$ is appropriate. On the other hand, if the objective is to ensure that the FPR is always lower than expected, then the $\chi^2(2)$ distribution should be preferred. Additionally, it is noteworthy that the Chi-square distribution is a special case of the Gamma distribution. Then, it would be possible to adapt the statistics of the deviance test by deriving a Gamma distribution with a specific shape parameter. Since this discrepancy appears when a single radionuclide needs to be tested against the Bkg-only hypothesis, the adaptation should depend only on the radionuclide library and the Bkg counting.

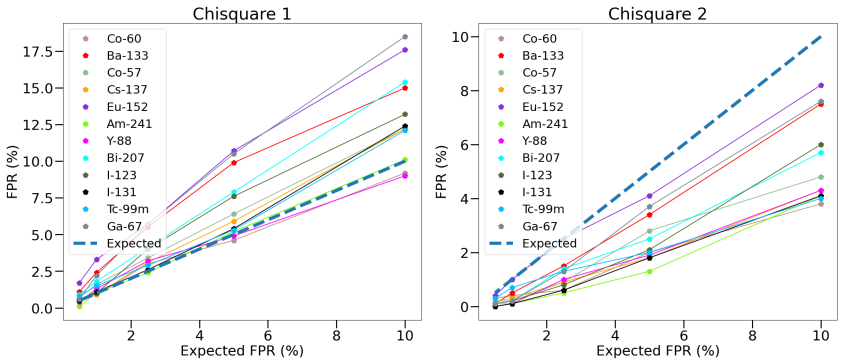

Fig. 4. FPR obtained using the deviance test based on $\chi^2(1)$ (left) and $\chi^2(2)$ (right) in the case where the radionuclide library contains only Bkg and one radionuclide. The average total counting is 1250. The x-axis represents the expected FPR used to define the threshold. The y-axis represents the FPR obtained by MoSeVa. Each curve corresponds to the result of the tested library containing Bkg and one radionuclide.

	Radionuclide											
Counting	^{60}Co	^{88}Y	^{241}Am	^{99m}Tc	^{131}I	^{137}Cs	^{57}Co	^{123}I	^{207}Bi	^{133}Ba	^{67}Ga	^{152}Eu
50000	5.6	7.3	4.6	8.3	8.7	6.9	6.1	5.6	8.8	10.0	8.7	9.8
5000	4.6	6.0	5.0	7.6	7.0	4.9	7.0	6.4	7.3	9.0	8.9	9.1
1250	4.6	4.9	5.2	5.3	5.4	5.9	6.4	7.6	7.9	9.9	10.5	10.7

Table 1: FPR obtained by means of the deviance test based on $\chi^2(1)$ with different counting of Bkg in the case where the tested library contains only Bkg and one radionuclide. The expected FPR is 5%. Each row corresponds to the result of each Bkg counting.

	Expected FPR(%)				
Counting	10	5	2.5	1	0.5
50000	14.4	9.1	5.9	1.9	1.2
5000	12.2	7.5	4.6	1.9	1.0
1250	12.3	5.9	2.8	1.5	0.8

Table 2: FPR obtained by means of the deviance test based on $\chi^2(1)$ with different expected FPR with the tested library of 12 radionuclides and Bkg. Each row corresponds to the result of each Bkg counting.

	Expected FPR(%)				
Counting	10	5	2.5	1	0.5
50000	5.9	2.3	1.2	0.6	0.2
5000	4.8	2.2	1.0	0.6	0.2
1250	2.8	1.7	0.9	0.2	0.2

Table 3: FPR obtained by means of the deviance test based on $\chi^2(2)$ with different expected FPR with the library of 12 radionuclides and Bkg. Each row corresponds to the result of each Bkg counting.

4.2.2. One or more radionuclides are present

This section evaluates the case where the source contains at least one radionuclide. Adding an extra radionuclide requires estimating an extra counting parameter, which would entail an extra number of degrees of freedom in the statistical deviance test. This would hold true if the latent variable λ were fixed. This is still a good approximation when λ is well estimated and does not vary significantly when adding an extra radionuclide. This is also the case when the estimators of both the counting a and the latent variable λ are equivalent to MLE and asymptotically follow Gaussian distributions. In that case, the estimator has the property of MLE and according to [16], the deviance asymptotically follows a $\chi^2(1)$ distribution.

The following experiment was carried out to illustrate this point. A mixture of Bkg with four radionuclides ^{60}Co, ^{137}Cs, ^{88}Y, ^{99m}Tc covering a large energy range (up to 2 MeV) was studied. The respective mixing weights were 0.5, 0.15, 0.1, 0.15, 0.1. The algorithm was evaluated for three cases: high, medium and low statistics with an average total counting of 100000, 10000 and 2500, respectively. In this study, Bkg is always present in the observed spectrum with a significant contribution, making the radionuclide identification at low statistics extremely complicated. Three steel sphere thicknesses were tested: 0.1, 3.6 and 20.1 mm, which cover different spectral signatures. The theoretical spectral signatures with a thickness of 20.1 mm and an example of simulated spectra at low statistics are given in Fig. 5. An example of the MoSeVa process is shown in Fig. 6.

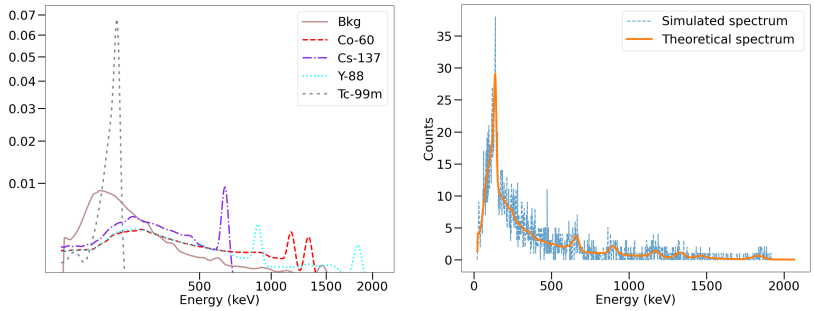

Fig. 5. Spectral signatures with 20.1 mm thickness (left) and a simulated spectrum at low statistics (right).

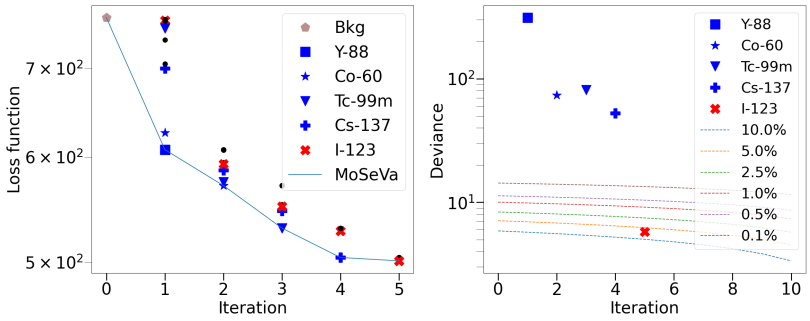

Fig. 6. Evolution of the loss function (left) and deviance (right) during iteration for the MoSeVa algorithm. Initially, only Bkg is imposed as present. Each black point in the figure on the left represents the loss function for the radionuclide not present in the source. At the first iteration, the algorithm selects the radionuclide that gives the smallest loss function: ^{88}Y is identified. The figure on the right shows the statistical deviance test for this radionuclide. The different-colored curve represents the threshold based on the χ^2 of a predefined FPR. As the deviance of this radionuclide is greater than the threshold, ^{88}Y is identified as present in the source. The algorithm then continues, and other radionuclides are selected.

Table 4 and Table 5 indicate the FPR obtained by P-OMP and MoSeVa with different average total counting and three different thicknesses. Generally, the results of MoSeVa are close to those of P-OMP (assuming the spectral signatures are known and fixed) and do not differ from the expected FPR (1%). It should be noted that P-OMP was demonstrated in [8] to provide an FPR that is almost identical to the expected value. In this present work, the FPR obtained by P-OMP is slightly different from the expected value because the number of Monte Carlo simulations (1000) is not sufficiently high. In most cases, no false negative is obtained by MoSeVa, except when the thickness is 20.1 mm with an average total counting of 2500. In this case, the FN almost comes from ^{137}Cs (only one FN comes from ^{99m}Tc). Since ^{137}Cs counting is very low (250), in some cases, the algorithm does not detect ^{137}Cs from the noise that originates from the Bkg.

The difficulty in detecting ^{137}Cs only arises at large thicknesses (*e.g.*, 20.1 mm) since the higher the thickness, the higher the effect of Compton scattering and the lower the amplitude of the ^{137}Cs γ-photon peak.

The FPR and FNR obtained at low statistics for a thickness of 20.1 mm are shown in Table 6. When the expected FPR imposed in the statistical test is reduced, the FPR obtained decreases, but the FNR obtained increases. Consequently, choosing a very low FPR (0.1%) can lead to a high FNR at low statistics. The selection of the expected FPR depends on the ratio to be controlled in priority (FPR or FNR) in decision-making.

The distributions of the estimated λ corresponding to the first anchor point for a 20.1 mm thickness at medium and low statistics are given in Fig. 7. These figures show that the distribution of λ is very close to the Gaussian distribution. Therefore, in this case, applying $\chi^2(1)$ as a stopping criterion can accurately control the FPR. In another case where the assumptions of applying $\chi^2(1)$ are not necessarily true, different mixing scenarios were studied in Appendix C. Even in these cases, the FPR obtained by MoSeVa is still close to the expected value.

	Counting		
Thickness	100 000	10 000	2500
0.1 mm	1.6	1.0	0.6
3.6 mm	1.6	1.3	1.1
20.1 mm	1.2	1.3	0.7
			0.6 | 0.1

Table 4: P-OMP: false prediction rate with different thicknesses and average total counting. The expected FPR is 1%. Values in black indicate no false negatives; false predictions are only false positives. Values in red mean the presence of false positives and false negatives. Values below these represent the FPR and FNR, respectively. For example, 0.7 is the false prediction rate, 0.6 is the FPR and 0.1 is the FNR.

	Counting		
Thickness	100 000	10 000	2500
0.1 mm	1.1	1.0	0.6
3.6 mm	1.7	1.1	1.6
20.1 mm	0.9	1.3	1.7
			1.2 | 0.6

Table 5: MoSeVa: false prediction rate with different thicknesses and average total counting. The expected FPR is 1%. Values in black indicate no false negatives; false predictions are only false positives. Values in red mean the presence of false positives and false negatives. Values below these represent the FPR and FNR, respectively. For example, 1.7 is the false prediction rate, 1.2 is the FPR and 0.6 is the FNR.

	Expected FPR(%)				
	5	2.5	1	0.5	0.1
FPR	4.5	2.3	1.2	0.6	0.3
FNR	0.2	0.4	0.6	0.9	1.9
False prediction rate	4.7	2.7	1.7	1.4	2.1

Table 6: FPR and FNR obtained by MoSeVa with different expected FPR at low statistics (2500) for a 20.1 mm thickness.

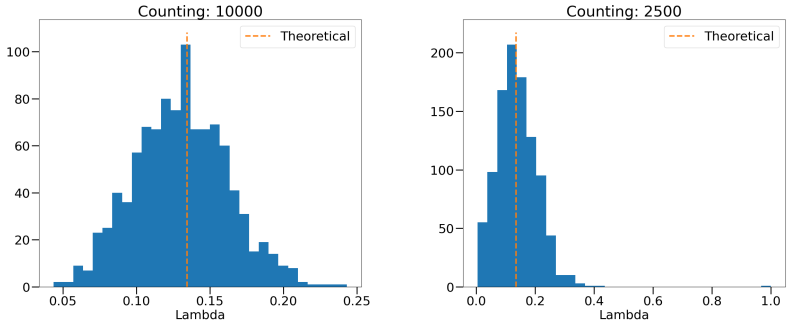

Fig. 7. The distribution of estimated λ of the first anchor point for a 20.1 mm thickness at medium statistics (left) and low statistics (right).

4.3. Investigations of identification performances for complex mixing scenarios

This section describes various complex mixing scenarios, including radionuclides with close γ-photon energies leading to larger correlations between their spectral signatures and potential confusion in the identification process. The correlation (as measured by the cosine similarity $cos(x, y) = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2}\sqrt{\sum y_i^2}}$) between the spectral signatures of all radionuclides for the 0.001 mm and 20.1 mm thickness are displayed in Fig. 8. At high thicknesses, low-energy peaks are absorbed by attenuation, leading to a higher correlation between spectral signatures than at low thicknesses.

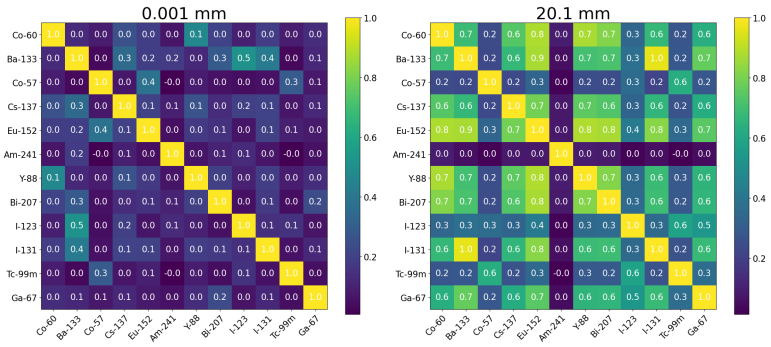

Fig. 8. The correlation matrix between spectral signatures for the 0.001 mm thickness (left) and 20.1 mm (right).

4.3.1. Mis-identification due to highly correlated spectral signatures

Owing to the sequential nature of the OMP algorithm and potential correlations between spectral signatures of the radionuclides to be tested, mis-identification may occur. This is particularly true for radionuclides that have a large number of γ-photon peaks such as ^{152}Eu, ^{133}Ba, ^{131}I. At the end of the selection process, the counting of the mis-identified radionuclides vanishes, but they are still present in the set of selected radionuclides.

A mixture of ^{137}Cs, ^{133}Ba with a thickness of 20.1 mm is studied to illustrate this problem. As observed in Fig. 9, the spectral signatures of ^{131}I and ^{133}Ba are highly correlated for energies below 400 keV. In addition, both ^{137}Cs and ^{131}I have a peak around 660 keV. Therefore, in some cases, the MoSeVa algorithm can mis-identify ^{131}I at the first iteration (Fig. 10) since ^{131}I better explains the observed spectrum than another radionuclides.

To overcome this phenomenon, an additional post-processing correction is proposed. When the selection procedure of MoSeVa is completed, an extra step is performed to look for the incorrectly detected radionuclide. Firstly, the algorithm selects the radionuclide that gives the least information about the observed spectrum. In other words, the algorithm selects the radionuclide so that the cost function is minimal without it. Then, the likelihood ratio test is performed to check the presence of this radionuclide. When there is no mis-identified radionuclide, the selected radionuclide is the same as the last iteration, so it always satisfies the statistical test. Hence, the result of the algorithm remains unchanged. In the case of mis-identification, the selected one is the incorrectly detected radionuclide, so it will almost be removed by the statistical test. Notably, the FPR obtained may be slightly higher than expected in this case, as several hypothesis tests are used. Fig. 11 displays the schema of this additional post-processing correction.

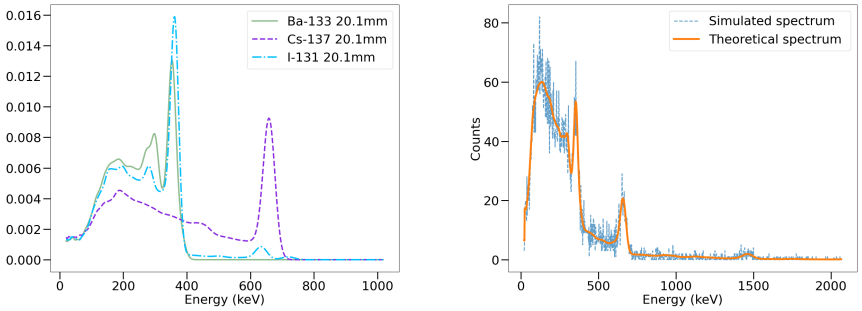

Fig. 9. Spectral signatures of ^{137}Cs, ^{133}Ba and ^{131}I for the 20.1 mm thickness (left) and a simulated γ-spectrum from ^{137}Cs and ^{133}Ba with the total counting 10000 (right). The spectral signatures of ^{131}I and ^{133}Ba are highly correlated for energies below 400 keV. The background is the same as that of the previous mixture.

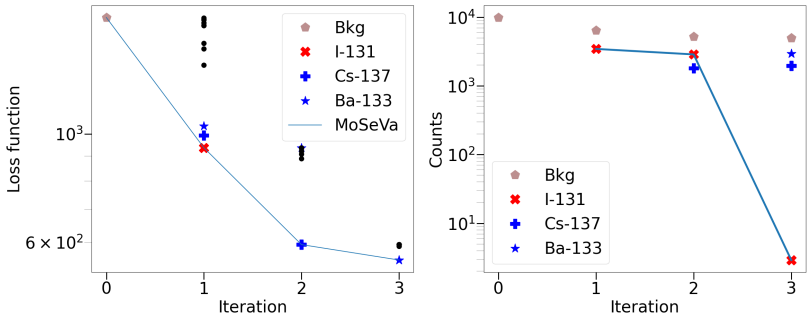

Fig. 10. Evolution of the loss function (left) and estimated counting (right) during iteration for the MoSeVa algorithm. Initially, only Bkg is imposed as present. Each black point in the figure on the left represents the loss function for the radionuclide not present in the radioactive source. At the first iteration, the algorithm selects the radionuclide that gives the smallest loss function: ^{131}I is identified. Then, the algorithm continues and other radionuclides: ^{137}Cs, ^{133}Ba are chosen. As a result, the radionuclides identified to be present in the source by MoSeVa are ^{131}I, ^{137}Cs, ^{133}Ba. The figure on the right shows the counting of selected radionuclides during each iteration. In the last iteration, the counting of ^{131}I decreases significantly and tends to be close to 0.

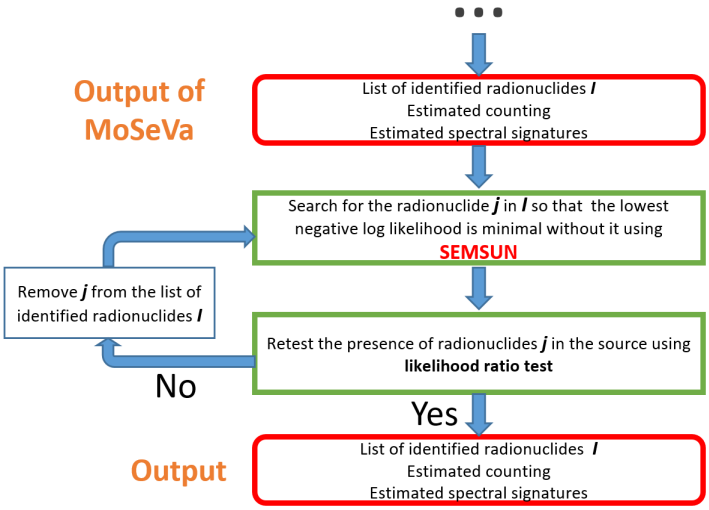

Fig. 11. Additional post-processing step for MoSeVa.

4.3.2. Identification performances for a mixture of ^{133}Ba, ^{137}Cs and Bkg

As shown above, a mixture of ^{133}Ba and ^{137}Cs can lead to a confusion with ^{131}I. The following numerical experiment was performed to evaluate the identification performances of the MoSeVa algorithm when the post-processing step was added. The mixture of Bkg with these two radionuclides was investigated for three cases: high, medium and low statistics. The corresponding mixing weights were 0.5, 0.3 and 0.2 for Bkg, ^{133}Ba and ^{137}Cs. As shown in Fig. 8, the spectral signatures of ^{131}I is highly correlated with ^{133}Ba. The spectral signatures of ^{133}Ba, ^{137}Cs and ^{131}I for a thickness of 0.1 mm and 20.1 mm are displayed in Fig. 12. The correlation between the spectral signatures of ^{131}I and ^{133}Ba is more significant at high thickness.

The FPR and FNR achieved by the P-OMP and modified MoSeVa are shown in Table 7 and Table 8. At high statistics, the resulting FPR is significantly higher than expected for thicknesses of 0.1 and 3.6 mm due to the IAE model. Since the spectral signatures are approximated by an IAE generative model, there is always a slight reconstruction error of the spectral signatures produced by the IAE model. Except for the high statistics case, the error that could originate from the IAE can be neglected since it is dominated by the Poisson noise. At high statistics, the impact of IAE error on the estimated spectral signature may become significant. In this case, the FPR is likely to increase. To verify this phenomenon, the same numerical evaluation was carried out using the best estimation of the theoretical spectral signatures by IAE $f(\lambda)$ instead of the theoretical values (X). With this modification, there is no approximation error, and the resulting FPR is close to the expected FPR (see Appendix D for more details). It should be noted that at high statistics, the estimated counting of the false positive radionuclide is very low compared to present radionuclides. It would then be possible to overcome this issue by further employing extra

constraints on the radionuclides counting based on some application-dependent prior knowledge.

At low statistics for a thickness of 20.1 mm, the FPR and FNR are high due to confusion between ^{133}Ba and ^{131}I. The false negative in this case differs from that of the ^{137}Cs and ^{99m}Tc discussed above. At low statistics, the spectral signatures of these two radionuclides with Poisson noise are similar, so in some cases, the cost function using ^{131}I is smaller than that using ^{133}Ba. In a nutshell, ^{131}I explains the simulated spectrum better than ^{133}Ba for some Poisson simulations and therefore the algorithm selects ^{131}I instead of ^{133}Ba. Ignoring the confusion between ^{131}I and ^{133}Ba, the difference between FNR and FPR (10.2 - 9.4) is exactly FPR in this case and close to the expected FPR (1%).

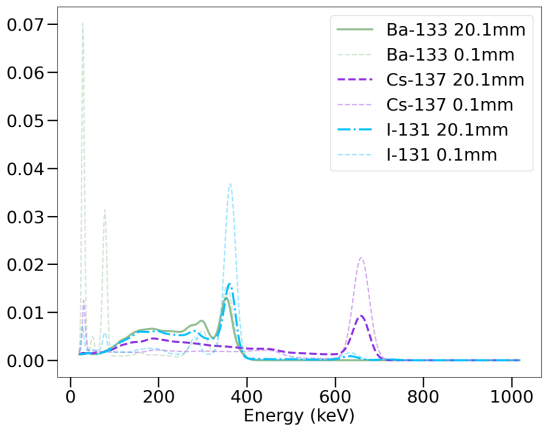

Fig. 12. Spectral signatures of ^{137}Cs, ^{133}Ba with 0.1 mm and 20.1 mm thicknesses.

	Counting		
Thickness	100 000	10 000	2500
0.1 mm	1.0	0.6	1.1
3.6 mm	1.2	0.9	1.2
20.1 mm	0.8	1.7	2.4
			2.4 |1.5

Table 7: P-OMP: false prediction rate with different thicknesses and average total counting. The expected FPR is 1%. Values in black indicate no false negatives; the false predictions are only false positives. Values in red mean the presence of false positives and false negatives. Values below these explain the FPR and FNR, respectively.

	Counting		
Thickness	100 000	10 000	2500
0.1 mm	2.5	0.9	0.8
3.6 mm	2.2	1.1	0.7
20.1 mm	1.3	1.9	10.2
			10.2 |9.4

Table 8: MoSeVa: false prediction rate with different thicknesses and average total counting. The expected FPR is 1%. Values in black indicate no false negatives; the false predictions are only false positives. Values in red mean the presence of false positives and false negatives. Values below these explain the FPR and FNR, respectively.

4.3.3. Identification performances for a mixture of ^{60}Co, ^{133}Ba, ^{207}Bi and Bkg

A more complex mixture of Bkg and three radionuclides ^{60}Co, ^{133}Ba, ^{207}Bi was studied with the corresponding mixing weights 0.4, 0.2, 0.2, 0.2. Fig. 13 displays the spectral signatures of these radionuclides for a thickness of 20.1 mm and an example of simulated γ-spectra. With this mixture, ^{152}Eu can be identified by the algorithm in some cases since it is a multi-γ emitter with photon energies close to those emitted by the other radionuclides. In addition, detecting ^{133}Ba can be difficult due to the correlation with the spectral signature of ^{131}I. As a result, radionuclide identification is even more complicated compared to the previous case.

The FPR and FNR obtained by P-OMP and MoSeVa are presented in Table 9 and Table 10. Generally, the FPR obtained is close to the expected values, except in a few exceptional cases. At high statistics with a thickness of 0.1 mm, the FPR is significantly higher than expected due to the error of the IAE, as described in the experiment presented before (see Appendix D for more details). The FNR with a thickness of 20.1 mm is due to confusion between ^{133}Ba and ^{131}I. If this confusion is ignored, FPR in this case (FPR-FNR) is close to the expected FPR (1%). It should be noted that with an average total counting of 10000, the confusion in this mixture is present because the average counting of ^{133}Ba (2000) is lower than in the mixture investigated in the previous section (3000).

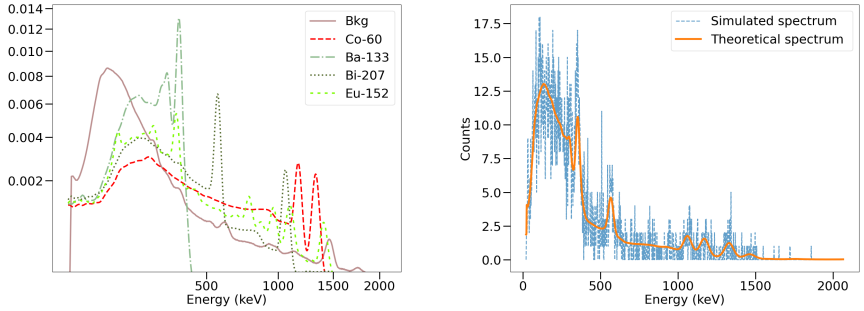

Fig. 13. Spectral signatures of ^{60}Co, ^{133}Ba, ^{207}Bi with the 20.1 mm thickness (left) and a γ-spectrum simulated from these radionuclides (right) at low statistics (2500).

	Counting		
Thickness	100 000	10 000	2500
0.1 mm	0.5	1.2	1.0
3.6 mm	0.6	1.2	1.1
20.1 mm	0.7	1.1	10.9
			10.9 | 10.0

Table 9: P-OMP: false prediction rate for different thicknesses and average total counting. The expected FPR is 1%. Values in black indicate no false negatives; the false predictions are only false positives. Values in red mean the presence of false positives and false negatives. Values below these explain the FPR and FNR, respectively.

	Counting		
Thickness	100 000	10 000	2500
0.1 mm	2.3	1.3	0.9
3.6 mm	1.3	0.9	1.2
20.1 mm	1.0	2.0	26.0
		2.0 |0.7	26.0 |25.2

Table 10: MoSeVa: false prediction rate for different thicknesses and average total counting. The expected FPR is 1%. Values in black indicate no false negatives; the false predictions are only false positives. Values in red mean the presence of false positives and false negatives. Values below these explain the FPR and FNR, respectively.

4.4. Quantification results of MoSeVa

The advantage of MoSeVa is that this algorithm can identify radionuclides and quantify their counting and spectral signatures simultaneously. In this section, the quantification result of a mixture of ^{60}Co, ^{137}Cs, ^{88}Y, ^{99m}Tc and Bkg was investigated. This algorithm is compared with Oracle, P-OMP and SEMSUN. Table 11 summarizes these comparisons.

	Oracle	P-OMP	SEMSUN	MoSeVa
Known present radionuclides	✓	✗	✓	✗
Known spectral signatures	✓	✓	✗	✗

Table 11: Characteristics of the different algorithms.

The spectral signatures estimated by MoSeVa (Fig. 14) are accurate compared to the theoretical value even at low statistics. Fig. 15 displays the relative errors of the estimated counting of four radionuclides for all algorithms at low statistics (2500) for the 20.1 mm thickness. Table 12 provides more details on the 25th, 75th percentile and median of the relative error of estimated counting. The MoSeVa result is similar to other methods, notably the SEMSUN algorithm. Indeed, the FPR obtained is small (1%), and even when there are false positives, the estimated counting of present radionuclides is not significantly different from the result when there are no false positives. Consequently, the result of MoSeVa is the same as that of SEMSUN. Regarding the impact of undetected radionuclides, some false negatives of ^{137}Cs and ^{99m}Tc are obtained, leading to an incorrect counting estimation (the relative error is -100%) for these particular cases, but the rest is similar to the SEMSUN result. These false negatives arise from the fact that the expected counting of these radionuclides is very low (250), so there is a certain probability that these radionuclides may not be detected. Compared to Oracle's result, MoSeVa's result is very close. In other words, the counting estimated by this algorithm is very similar to the case where spectral signatures and present radionuclides are known.

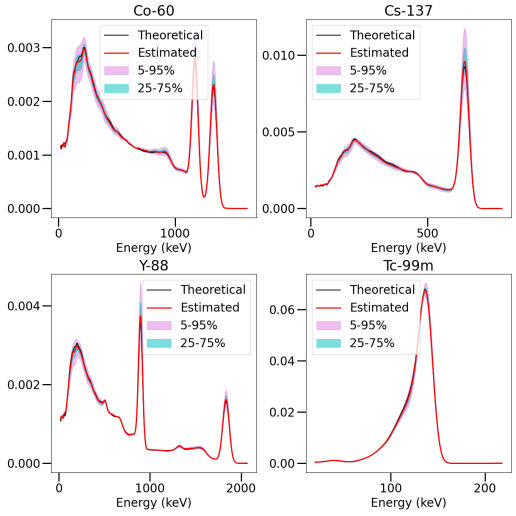

Fig. 14. Estimated spectral signature of 4 radionuclides at low statistics (2500) with a thickness of 20.1 mm. Theoretical represents theoretical spectral signatures, and Estimated represents the median of the estimated spectral signatures for all Monte Carlo simulated γ-spectra. 25-75 % represents the interval between the 25 and 75 percentile. Same for 5-95%.

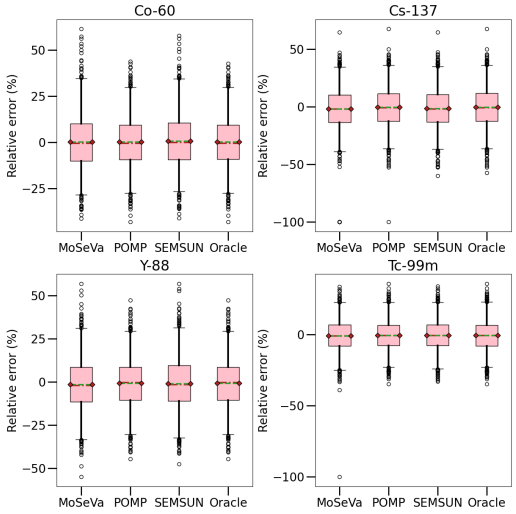

Fig. 15. Low statistics: relative errors of estimated counting for four radionuclides. At the center of the boxplot, the red line represents the median, and the blue line represents the mean. The box is composed of the first and third quartiles. The whiskers represent the 2nd and 98th percentiles of the data.

	Bkg	^{60}Co	^{137}Cs	^{88}Y	^{99m}Tc
MoSeVa	**0.4**[-4.1, 5.6]	**-0.3**[-10.0, 10.1]	**-1.8**[-13.7, 10.6]	**-2.0**[-11.5, 8.8]	**-0.8**[-8.1, 6.7]
P-OMP	**0.1**[-4.0, 4.2]	**-0.5**[-9.4, 9.5]	**-0.5**[-12.5, 11.6]	**-0.4**[-10.7, 8.5]	**-0.7**[-7.8, 6.6]
SEMSUN	**0.2**[-4.3, 5.2]	**0.0**[-9.4, 10.6]	**-1.5**[-13.3, 11.0]	**-1.5**[-11.1, 9.5]	**-0.5**[-7.8, 6.9]
Oracle	**0.1**[-3.9, 4.2]	**-0.5**[-9.2, 9.5]	**-0.5**[-12.5, 11.8]	**-0.4**[-10.7, 8.5]	**-0.7**[-7.9, 6.6]

Table 12: Low statistics: relative errors of estimated counting (%) of all algorithms for present radionuclides. The values in each cell are the median, 25th percentile and 75th percentile, respectively.

5. Conclusion

A novel hybrid unmixing algorithm, MoSeVa, has been developed to automatically identify and quantify γ-emitting radionuclides, taking into account spectral variability that originates from the measurement environment of the radioactive source to be analyzed. MoSeVa combines a variability-aware spectral unmixing algorithm, coined SEMSUN, and a model selection technique P-OMP. SEMSUN builds upon a pre-trained machine learning model to capture spectral deformations. The model selection procedure relies on a sequential multiple single-hypothesis testing process, with a stopping procedure based on the LRT. Adaptations are introduced to account for the spectral variability accurately. First, the LRT depends on the number of degrees of freedom of the Chi-squared distribution of the statistical deviance, which can be modified according to the expected false positive rate. Second, an additional post-processing correction

is proposed to eliminate mis-identification when the spectral signatures of radionuclides are highly correlated. The algorithm has been evaluated in different mixing scenarios with different statistics and a library of 12 radionuclides. In most cases, the false positive rates obtained by this algorithm are close to the result given when the spectral signatures are known and also close to the expected value (1%). When no radionuclide is present in the radioactive source, different deviance distributions can be used, depending on the specific requirements of the application. At high statistics, the reconstruction error of spectral signatures of IAE can lead to increased false positive rates in some cases but can be reduced when applied with additional prior knowledge of radionuclide counting. Concerning quantification results, the distributions of estimated counting of all radionuclides are similar to those obtained when spectral signatures and identified radionuclides are known. This study demonstrates the ability of the hybrid unmixing approach to handle spectral variability for the automatic identification and quantification of γ-emitting radionuclides. While the evaluation of this work was carried out using simulations, future research will evaluate this approach using experimental measurements.

Appendix A. Pseudocode of MoSeVa

Input:
Observed spectrum y and normalized natural background X_{Bkg}
Pretrained IAE model $\mathcal{M}$ for modelling the spectral signatures of N radionuclides
Initialization:
List of identified radionuclides: $I_0 = [0]$ (Bkg)
List of radionuclides to be tested: $I = [1..N]$
The mask for the list of radionuclide: $m = [m_1, .., m_N] = [0, .., 0]$
The cost function $L_0 = D(y||X_{Bkg} \sum(y))$
Set $Flag = 1$
while *Flag=1 and* $I \neq \emptyset$ **do**
  for $i \in I$ **do**
    Add a radionuclide to a list of identified radionuclides $I_{test} \leftarrow I_0 \cup j$
    Update the value of mask of the radionuclide in I_{test} : $m[I_{test}] = 1$
    Compute the estimated λ, counting and cost function :
    $\lambda^{(i)}, a^{(i)}, L^{(i)}_{test} = SEMSUN(\mathcal{M}, y, X_{Bkg}, \lambda, m)$
  end
  Select the radionuclide which has the lowest cost function: $j = argmin_i L^i_{test}$
  Define the decision threshold:
  if *Only Bkg in the list of identified radionuclide:* $I_0 = [0]$ **then**
    $DT = z_{\chi^2(1)}(1 - 2\alpha/dim(I))$ or $DT = z_{\chi^2(2)}(1 - 2\alpha/dim(I))$
  else
    $DT = z_{\chi^2(1)}(1 - 2\alpha/dim(I))$
  Test the deviance:
  if $-2(L^{(j)}_{test} - L0) > DT$ **then**
    Update the list of identified radionuclide: $I0 \leftarrow I0 \cup j$
    Remove this radionuclide from the list that needs to be tested: $I \leftarrow I \setminus j$
    Update the cost function, λ, a: $L_0 \leftarrow L^{(j)}_{test}$, $\lambda \leftarrow \lambda^{(j)}$, $a \leftarrow a^{(j)}$
  else
    Set Flag=0
end
Additional tests:
Set $Flag = 1$
while *Flag=1 and* $dim(I0) > 2$ **do**
  for $i \in I0$ **do**
    Remove a radionuclide from a list of identified radionuclides $I_{test} \leftarrow I_0 \setminus i$
    Update the value of mask of the radionuclide in I_{test} : $m[I_{test}] = 1$
    Compute the estimated λ, counting and cost function :
    $\lambda^{(i)}, a^{(i)}, L^{(i)}_{test} = SEMSUN(\mathcal{M}, y, X_{Bkg}, \lambda, m)$
  end
  Select the radionuclide which has the lowest cost function: $j = argmin_i L^i_{test}$
  Define the decision threshold: $DT = z_{\chi^2(1)}(1 - 2\alpha/(dim(I) + 1))$
  Test the deviance:
  if $2(L^{(j)}_{test} - L0) > DT$ **then**
    Update the list of identified radionuclides: $I0 \leftarrow I0 \setminus j$
    Add this radionuclide to the list of non-present radionuclides: $I \leftarrow I \cup j$
    Update the cost function, latent variable, counting: $L_0 \leftarrow L^{(j)}_{test}$, $\lambda \leftarrow \lambda^{(j)}$,
    $a \leftarrow a^{(j)}$
  else
    Set Flag=0
end
Output:
The list of identified radionuclides $I0$
Estimated latent variable λ
Estimated counting a

Algorithm 1: Pseudocode MoSeVa

Appendix B. Applying $\chi^2(1)$ leads to a higher-than-expected FPR when no radionuclide is present

Let $\hat{a}_{Bkg}^{(1)}$ be the estimated Bkg counting when there is only Bkg in the list of identified radionuclides $I0$. $\hat{a}_{Bkg}^{(2)}$, $\hat{\lambda}^{(2)}$, $[f(\hat{\lambda}^{(2)})]_i$, $\hat{a}_i^{(2)}$ are respectively the estimated Bkg counting, the estimated λ, the estimated spectral signature of the radionuclide concerned and its estimated counting for the mixture of Bkg and radionuclide i. $\bar{a}_{Bkg}^{(1)}$, $\bar{a}_i$, $\bar{X}_i$ are respectively the estimated Bkg counting, the estimated counting and the spectral signature of radionuclide i generated from the IAE model for the mixture of Bkg and radionuclide i when the spectral signatures are fixed. Let D_1 be the deviance between the Bkg and the mixture of Bkg and a known fixed spectral signature of radionuclide i and D_2 the deviance between the Bkg and the case where the spectral signature of radionuclide i must be estimated.

$$D_1 = 2(D(y||\hat{a}_{Bkg}^{(1)}X_{Bkg}) - D(y||\bar{a}_{Bkg}^{(1)}X_{Bkg} + \bar{a}_i\bar{X}_i))$$

$$D_2 = 2(D(y||\hat{a}_{Bkg}^{(1)}X_{Bkg}) - D(y||\hat{a}_{Bkg}^{(2)}X_{Bkg} + \hat{a}_1[f(\hat{\lambda}^{(2)})]_i))$$

Let $z_{1-\alpha}$ be $1-\alpha$ quantile of $\chi^2(1)$, so $P(D_1 > z_{1-\alpha}) = \alpha$. The cost function when estimating the spectral signature is always less than or equal to the cost function when the spectral signature is fixed, so $D_2 \geq D_1$. As a result, $P(D_2 > z_{1-\alpha}) \geq P(D_1 > z_{1-\alpha}) = \alpha$. This can be interpreted as the fact that by applying the $\chi^2(1)$, the FPR in this case is higher than expected.

Appendix C. Assumption for applying $\chi^2(1)$ when at least one radionuclide is present

When at least one radionuclide is present in the radioactive source, an important assumption is made to use $\chi^2(1)$: λ is well estimated and does not vary significantly when an extra radionuclide is added or asymptotically follows the Gaussian distribution. However, λ belongs to [0,1]$\times$ [0,1] due to the simplex and positivity constraint. When the theoretical spectral signature is close to the anchor point, the estimated λ distribution may not be Gaussian. To assess this problem, the same numerical evaluation 4.2.2 was carried out using the theoretical thickness close to the thickness of the anchor points (0.001 and 30 mm). At low statistics, a histogram of the estimated λ with the 30 mm thickness is presented in Fig. C.16. In this case, the estimated λ does not follow the Gaussian distribution.

The FPR and FNR obtained by MoSeVa are presented in Table C.14. At high statistics, with a thickness of 0.001 mm and 30 mm, the FPR obtained is larger than expected due to the reconstruction error of the IAE model. With a total counting of 10000, the FPR obtained is not significantly different from that of expected. The low-statistics result is consistent with those obtained in the previous section: the false negatives appear and almost come from ^{137}Cs. Even though the λ distribution is not Gaussian, the identification result is very

close to that expected when using $\chi^2(1)$. This can be explained by the fact that when a non-present radionuclide is added to the true model, the estimated spectral signatures or λ remain almost unchanged. Fig. C.17 illustrates the λ difference and the NMSE of the spectral signatures between these models. The λ changes are relatively minor in this case, which means that the estimated spectral signatures also change relatively little (high NMSE). As a result, when a non-present radionuclide is added, the algorithm behaves as in the case where spectral signatures are fixed, and the LRT test with $\chi^2(1)$ is still valid.

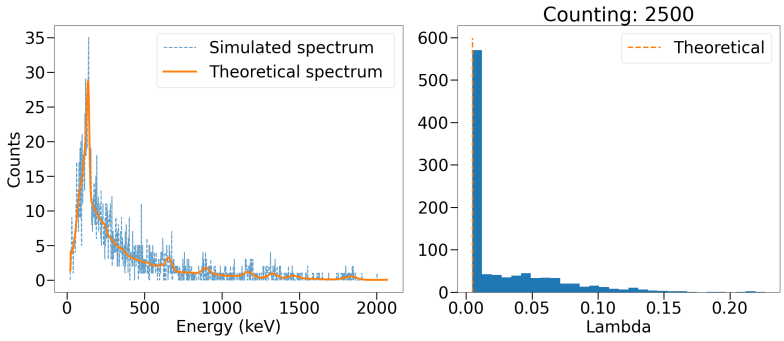

Fig. C.16. A simulation of a mixture of ^{60}Co, ^{137}Cs, ^{88}Y and ^{99m}Tc with the thickness of 30 mm at low statistics (left) and estimated λ of the first anchor point (right).

	Counting		
Thickness	100 000	10 000	2500
0.001 mm	1.4	0.8	1.2
28 mm	0.8	1.0	1.6
			0.9 |0.7
30 mm	0.8	1.1	1.8
			0.9 |1.0

Table C.13: False prediction rate with different thicknesses and average total counting. Each cell represents the result for MoSeVa and P-OMP, respectively. The expected FPR is 1%. Values in black indicate no false negatives; the false predictions are only false positives. Values in red indicate false positives and negatives in the results obtained by the algorithms. Values below these explain the FPR and FNR, respectively.

	Counting		
Thickness	100 000	10 000	2500
0.001 mm	2.9	1.3	1.4
28 mm	0.9	1.2	1.7 0.8 |0.9
30 mm	1.8	1.3	2.0 1.1 |0.9

Table C.14: False prediction rate with different thicknesses and average total counting. Each cell represents the result for MoSeVa and P-OMP, respectively. The expected FPR is 1%. Values in black indicate no false negatives; the false predictions are only false positives. Values in red indicate false positives and negatives in the results obtained by the algorithms. Values below these explain the FPR and FNR, respectively.

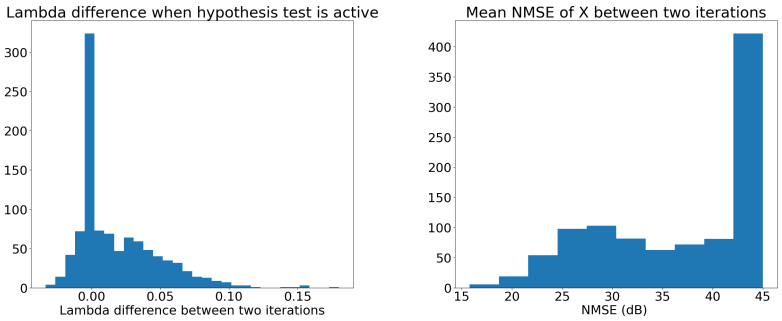

Fig. C.17. Left: λ difference between two models: one is the true model (contains all present radionuclides), and the other is the true model plus an extra false negative (non-present radionuclide) at low statistics and a thickness of 30 mm. Right: mean NMSE of spectral signatures of identified radionuclides between these two models.

Appendix D. Impact of IAE error on identification performance

At high statistics, the FPR obtained is significantly higher than expected due to the reconstruction error of spectral signatures of the IAE model. To verify this, the same numerical evaluation was carried out using the IAE's best approximations of theoretical spectral signatures instead of the theoretical values. D.15 and D.16 show the result of this comparison for the mixtures discussed above. When IAE-estimated spectral signatures are used, no IAE error is present in the algorithm, so the resulting FPR is close to the expected FPR.

	Thickness		
	0.1 mm	3.6 mm	20.1 mm
Geant4	2.5	2.2	1.3
IAE	1.1	1.2	0.7

Table D.15: False prediction rate of MoSeVa with different thicknesses at high statistics for the mixture of ^{133}Ba, ^{137}Cs and Bkg. Geant4 represents the case where theoretical spectral signatures are used. IAE represents the case where the best estimations of theoretical spectral signatures by IAE are used.

	Thickness		
	0.1 mm	3.6 mm	20.1 mm
Geant4	2.3	1.3	1.0
IAE	0.7	0.9	0.9

Table D.16: False prediction rate of MoSeVa with different thicknesses at high statistics for the mixture of ^{60}Co, ^{133}Ba, ^{207}Bi and Bkg. Geant4 represents the case where theoretical spectral signatures are used. IAE represents the case where the best estimations of theoretical spectral signatures by IAE are used.

References

[1] G. Gilmore, Practical gamma-ray spectroscopy, John Wiley & Sons, 2008.

[2] P. EURO, Combating illicit trafficking in nuclear and other radioactive material, IAEA Nuclear Security 6 (2007) 3–12.

[3] G. Daniel, F. Ceraudo, O. Limousin, D. Maier, A. Meuris, Automatic and real-time identification of radionuclides in gamma-ray spectra: a new method based on convolutional neural network trained with synthetic data set, IEEE Transactions on Nuclear Science 67 (4) (2020) 644–653.

[4] A. N. Turner, C. Wheldon, T. K. Wheldon, M. R. Gilbert, L. W. Packer, J. Burns, M. Freer, Convolutional neural networks for challenges in automated nuclide identification, Sensors 21 (15) (2021) 5238.

[5] S. Galib, P. Bhowmik, A. Avachat, H. Lee, A comparative study of machine learning methods for automated identification of radioisotopes using NaI gamma-ray spectra, Nuclear Engineering and Technology 53 (12) (2021) 4072–4079.

[6] W. Kim, K. Ko, J. Park, S. Lee, H. Yun, G. Cho, Deep learning-based gamma spectroscopic analysis considering multiple variables for in situ applications, Radiation Physics and Chemistry 226 (2025) 112261.

[7] B. Paleti, G. H. Sastry, Identification of gamma emitting natural isotopes in environmental sample spectra: convolutional neural network approach, Journal of Radioanalytical and Nuclear Chemistry 332 (12) (2023) 5273–5281.

[8] R. André, C. Bobin, J. Bobin, J. Xu, A. de Vismes Ott, Metrological approach of γ-emitting radionuclides identification at low statistics: application of sparse spectral unmixing to scintillation detectors, Metrologia 58 (1) (2021) 015011.

[9] J. Xu, J. Bobin, A. de Vismes Ott, C. Bobin, Sparse spectral unmixing for activity estimation in γ-ray spectrometry applied to environmental measurements, Applied Radiation and Isotopes 156 (2020) 108903.

[10] P. Malfrait, J. Bobin, A. de Vismes Ott, Spectral unmixing of multitemporal data in gamma-ray spectrometry, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 1045 (2023) 167547.

[11] P. Malfrait, J. Bobin, A. de Vismes Ott, Online spectral unmixing in gamma-ray spectrometry, Applied Radiation and Isotopes 201 (2023) 111011.

[12] M. S. Badawi, A numerical simulation method for calculation of linear attenuation coefficients of unidentified sample materials in routine gamma ray spectrometry, Nuclear Technology and Radiation Protection 30 (4) (2015) 249–259.

[13] M. I. Abbas, M. S. Badawi, A. A. Thabet, Y. N. Kopatch, I. N. Ruskov, D. Grozdanov, S. Noureddine, N. Fedorov, M. M. Gouda, C. Hramco, et al., Efficiency of a cubic NaI (Tl) detector with rectangular cavity using standard radioactive point sources placed at non-axial position, Applied Radiation and Isotopes 163 (2020) 109139.

[14] F. M. de Oliveira, G. Daniel, O. Limousin, Artificial gamma ray spectra simulation using Generative Adversarial Networks (GANs) and Supervised Generative Networks (SGNs), Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 1047 (2023) 167795.

[15] D. T. Phan, J. Bobin, C. Thiam, C. Bobin, A hybrid machine learning unmixing method for automatic analysis of γ-spectra with spectral variability, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 1060 (2024) 169028.

[16] S. S. Wilks, The large-sample distribution of the likelihood ratio for testing composite hypotheses, The annals of mathematical statistics 9 (1) (1938) 60–62.

[17] American national standard for performance criteria for spectroscopy-based portal monitors used in homeland security, ANSI N42.38-2015 (Revision of ANSI N42.38-2006) (2016) 1–59.

[18] F.-X. Dupé, S. Anthoine, A greedy approach to sparse poisson denoising, in: 2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP), 2013, pp. 1–6. doi:10.1109/MLSP.2013.6661993.

[19] A. DasGupta, Asymptotic theory of statistics and probability, Vol. 180, Springer, 2008.

[20] R. A. Armstrong, When to use the bonferroni correction, Ophthalmic and Physiological Optics 34 (5) (2014) 502–508.

[21] J. Bobin, R. C. Gertosio, C. Bobin, C. Thiam, An autoencoder-based model for learning regularizations in unmixing problems, Digital Signal Processing (2023) 104058.

[22] Y. Xu, W. Yin, A globally convergent algorithm for nonconvex optimization based on block coordinate update, Journal of Scientific Computing 72 (2) (2017) 700–734.

[23] D. C. Liu, J. Nocedal, On the limited memory BFGS method for large scale optimization, Mathematical programming 45 (1-3) (1989) 503–528.

[24] S. Agostinelli, J. Allison, K. a. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, et al., Geant4—a simulation toolkit, Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506 (3) (2003) 250–303.