CrossMark

Automatic identification and quantification of radionuclides in gamma spectra using numerical optimization

Kenneth Dayman[1] · Steven Biegalski[1]

Abstract We propose a novel method for gamma-ray spectral analysis that populates and solves an arbitrarily large system of radionuclides in order to explicitly model the production and loss of chains of radionuclides during counting. This solution is embedded into an optimization problem in order to resolve spectral interferences and find the nuclide mixture that best explains the measured spectrum. We demonstrate the efficacy of this method using a synthetic example with complex spectral interferences, short half lives, and nuclide transmutation, as well as a spectrum collected of a multi-gamma standard typically used for efficiency calibration.

Keywords Gamma-ray spectroscopy · Optimization · Data analysis

Introduction

Gamma-ray spectroscopy is a multi-elemental technique that can identify and quantify multiple gamma-emitting radionuclides within a sample. Unlike radiochemical separation and destructive analysis, gamma-ray spectroscopy does not require pre-chemistry prior to analysis. The elimination of additional chemistry potentially allows non-destructive analysis to be conducted in the field using portable spectrometers. In addition, eliminating pre-chemistry removes potential biases that may arise due to chemical fractionation and incomplete separations.

Correctly accounting for these phenomena typically requires preparing and analyzing additional samples such as spiked standards.

Current commercially available gamma-ray spectroscopy software is not able to handle the nuclide identification and quantification tasks that would be required for post-detonation nuclear debris samples that have many short-lived nuclides that are present along with their parent and daughter products. Radionuclides typically emit multiple gamma rays during decay and de-excitation. As more gamma-emitting nuclides are present in the sample, increasingly complex spectra that feature tens or hundreds of distinct photopeaks are obtained. Each photopeak may overlap with other peaks—these multiplets must be fitted and attributed to a collection of radionuclides in order to resolve the interference and obtain accurate activity measurements for each contributing radionuclide.

The following section presents a summary of the physics of gamma spectroscopy, typical spectral features, and the main objectives of spectral analysis. Section 3 gives a description of a new method for gamma-ray spectral analysis. Sections 4 and 5 describe the experiments used to test this new methodology and summarize the results.

Background

During the decay of a gamma-emitting radionuclide, characteristic photons are emitted at known rates. These rates are given by the gamma intensities of each photon. The energy of these photons identify the radionuclide and the count rate gives the activity (more specifically, the number of decays during counting). Multiple nuclides may emit photons of a particular energy; however, the relative

✉ Kenneth Dayman
  kenneth.dayman@gmail.com

[1] University of Texas at Austin, 10100 Burnet Rd., Bldg. 159, Austin, TX 78758, USA

Springer

emission of the characteristic photons form a unique multivariate signature for the radionuclide.

A sample with many gamma-emitting radionuclides such as fission products will produce complex gamma spectrum with tens or hundreds of distinct photopeaks and exhibit a number of spectral features that challenge traditional analysis. It is almost guaranteed that multiple nuclides will contribute to many of the photopeaks. This effect is known as spectral interference. Typically, the analyst will find a second characteristic photopeak that does not have an interference and use the area of this second peak along with the gamma intensity of each photon to estimate the analyte's contribution to the questionable peak. In principle, if there does not exist an interference-free peak, then the interference on the second peak may be resolved and then the analysis repeated recursively repeated until all interferences are resolved; however, this proceed is not practical when the number of interferences becomes large [1]. This effect compounds the complexity of nuclide identification, as spectral interference resolution is convoluted with hypothesis testing to determine the sources of the observed photopeaks.

When gamma emission rates are sufficiently high (relative to the pulse-processing/resolving time of the spectrometer), sum peaks may be observed in the spectrum. Photons that are summed produce a new photopeak in the spectrum that is not characteristic of their source[1] and reduce the observed counts in the source's characteristic photons. These two are sometimes called summing-in and summing-out effects. These effects may be corrected much in the same way spectral interferences are accounts for; however, the problem of complexity remains the same.

In addition, photons with energy of at least 1.022 MeV may undergo pair-production within the detector, producing a positron, which annihilates and leads to 0.511 MeV peaks in the spectrum, as well as peaks 0.511 or 1.022 MeV below the original energy of the photon. These are called single- and double-escape peaks, respectively.

Theory

To overcome the shortcomings of traditional gamma-ray spectra analysis, our methodology combines several parts and innovations. First, an eye-readable and easily modified library has been developed from ENSDF, unlike existing data repositories such as ENDF [2]. At runtime, this library is used in performing a peak search that identities potential gamma-emitting radionuclides within

the sample, populating a list of hypothesized nuclides and decay products, and setting up the mathematical problem, which we develop below.

First, consider a decay chain of radionuclides as shown in Fig. 1. Five nuclides are shown and the arrows show how the nuclides decay into one another. Note, that nuclide N_1 is stable and nuclide N_2 decays to N_1 and N_4 with some branching ratio.

For the most accurate determination of the activities of nuclides 2,3,4, and 5, their activities must be found simultaneously, i.e., by solving an appropriate system of equations. In general, for a set of M radionuclides, these equations have the form

$$\frac{\mathrm{d}}{\mathrm{d}t}(N_i) = \sum_j \lambda_{i-j} N_{i-j} - \lambda_i N_i, \quad i = 1, 2 \ldots M, \tag{1}$$

where N_i is the time-dependent number density of the ith nuclide, λ_i is the decay constant of the ith nuclide, λ_{i-j} is the partial decay constant of the $i-j$ nuclide to the ith nuclide (i.e., λ_{i-j} times the branching ratio that populates nuclide i). Note, daughter products are given higher indices than parent nuclides, thus the summation over index j denotes sources from all possible parent nuclides. The system of equations shown in Eq. (1) are called the Bateman equations.

By collecting each of the M equations shown in Eq. (1) into vector $N \in \mathbb{R}^M$ and matrix $\mathbf{B} \in \mathbb{R}^{M \times M}$, the above system may be transformed into a straightforward matrix formulation:

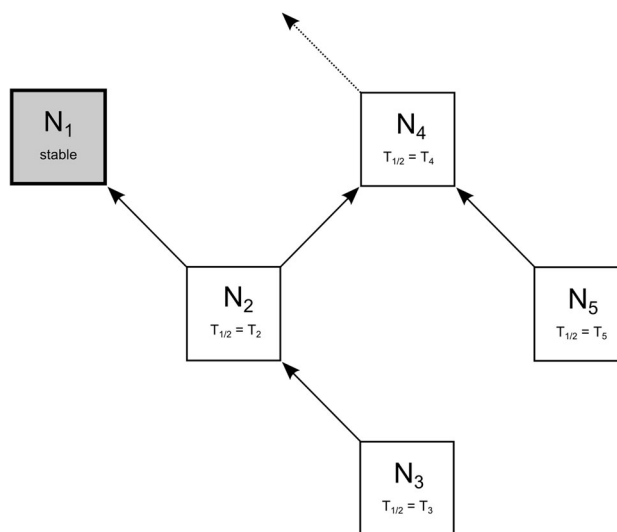$$\frac{\mathrm{d}}{\mathrm{d}t}(N) = \mathbf{B}N. \tag{2}$$

Fig. 1 A generic decay chain of four radionculides. In order to correctly quantify each nuclide, decay during counting and transmutation from 5 to 4, 3 to 2, and 2 to 4 must be taken into account

---

[1] If there is peak at or very near the energy of the sum peak, then the count rate will be inflated, likely leading to an overestimation of activity.

Note the matrix $\mathbf{B}$ is sparse and lower-triangular as daughter product(s) are given higher indices than parents.

The solution to this system is obtained using matrix exponentials, and multiplying through by the diagonal matrix of decay constants, λ, gives the corresponding activities, shown below.

$$A(t) = \Lambda\left(e^{(t\mathbf{B})}\right)N_0. \tag{3}$$

Given the lower-diagonal form for $\mathbf{B}$, the matrix $\mathbf{B}$ is non-singular[2], and the integral of Eq. (3) is easily computed in order to determine the total number of decays that occur during the counting time, $t \in [0, t_c]$:

$$D = \int_0^{t_c} \left(\Lambda\left(e^{(t\mathbf{B})}N_0\right)\right)\, \mathrm{dt} = \Lambda\mathbf{B}^{-1}\left(e^{(t_c\mathbf{B})} - \mathbf{I_M}\right)N_0, \tag{4}$$

where $\mathbf{I_M}$ is the $M \times M$-dimensional identity matrix. Using this equation, the number of decays each hypothesized nuclide undergoes during the measurement, D, is calculated given the initial values in the vector N_0.

Finally, the number of decays of each radionuclide may be used to predict the counts in each photopeak in a spectrum by multiplying D by a response matrix, $\mathbf{R} \in \mathbb{R}_+^{M \times K}$, defined element-wise as

$$R_{ij} = \gamma_j, \tag{5}$$

where γ_j is the gamma intensity for the jth photon coming from the ith hypothesized nuclide.

The initial quantities of each nuclide in the sample, n_0, may be found by predicting the counts in each photopeak observed in the spectrum, $\mathbf{R}^T D$, and comparing this vector to the efficiency-corrected counts in each peak observed in the spectrum: $s_i = c_i / \epsilon(E_i)$, where c_i is the counts of the ith photopeak and $\epsilon(E_i)$ is the associated efficiency. The weighted difference between the predicted and observed photopeak areas is the objective of a constrained optimization problem that is used solved to find the initial quantities of each nuclide in the sample.

$$\hat{N}_0 = \underset{N_0 \in \mathbb{R}_+^M}{\mathrm{argmin}} \left\| \mathbf{W}\left(s - \mathbf{R}^T\Lambda\mathbf{B}^{-1}\left(e^{(t_c\mathbf{B})} - \mathbf{I_M}\right)N_0\right)\right\|_1$$
$$\hat{A}_0 = \Lambda\hat{N}_0 \tag{6}$$

This convex optimization formulation seeks the optimal mixture of nuclides in the sample that explains the measured response as well as enforces correct time behavior if multiple spectra are used. The use of the L_1 norm gives preference to sparse solutions, i.e., solutions where as many nuclide concentrations are zero or near-zero as possible.

---

[2] A simple test is to compute the determinant. Since $\mathbf{B}$ is triangular, the determinant is equal to the trace of $\mathbf{B}$, $\mathrm{tr}(\mathbf{B}) = \sum_{i=1}^M \lambda_i$. This sum is nonzero, making the matrix nonsingular and invertible.

Due to the use of the L_1 norm, there is no closed-form solution for Eq. (6), and a solution is produced using the SeDuMi solver and the CVX Matlab interface [3–5]. In principle, there is no inherent limit to the size of the system that may be modeled with the above formulation.

Experiments

This section outlines two experiments testing the ability of the formulation summarized in the previous section to simultaneously identify and quantify radionuclides from a list of photopeak areas. Two experiments are presented: first, a synthetic example with measurement noise, spectral interference, and unobserved photopeaks; second, real data collected of a multi-gamma standard typically used for detector efficiency calibrations.

Synthetic example

In this example we test the method's ability to correctly resolve spectral interferences and quantify the initial activity of a decay chain of radionuclides. We have devised a scenario that features complex spectral interferences, decay chains, and short half lives relative to the counting time. The scenario features seven hypothesized nuclides contributing to seven photopeaks. This example does not utilize the peak search features described in the previous section—these features are utilized in the second and third examples.

Table 1 summarizes the scenario parameters. The most important feature of this dataset is the degree of spectral interference. Only nuclide A has an interference-free photopeak. Nuclides C and D each have a single characteristic photon, but they interfere with each other. The only way to resolve this interference is by enforcing the time behavior dictated by Eq. (1), which couples nuclides C and D to nuclides B and E, respectively. Each of the nuclides has a short half life relative to the counting time (see below), making correction for decay during counting necessary for correct activity determination.

The hypothetical nuclear data were used to calculate the expected number of counts in each photopeak following a count time of 600 s, and Gaussian noise was added to the results. Uncertainties in the peak areas were taken to be the square root of the area, which is consistent with radiation counting at moderate count rates. These values are given in Table 2.

Long-lived multi-gamma standard

In this example, we aim to demonstrate the method described in Sect. 3 can correctly identify gamma-emitting

Table 1 Input values used in the synthetic example

Nuclide	Activity (Bq)	Half life (s)	Daughter	peak energy (keV)	Intensity (%)	Efficiency (%)
A	1000	600	n/a	163.3	50	15
				276.5	60	10
				584.4	35	5
B	600	250	C	276.5	25	10
				412.7	15	7
				728.4	35	4
C	100	750	n/a	383.0	28.5	7.5
D	500	350	E	383.0	25	7.5
E	250	1000	n/a	584.4	50	5
				728.4	20	4
				649.0	85	4.5
F	600	450	n/a	276.5	40	10
				649.0	10	4.5
G	0	700	n/a	412.7	25	7

Quoted activities are at the beginning of counting. Nuclide B decays into nuclide C, and nuclide D decays into E. Nuclide G does not contribute to the peak areas (see Table 2)

Table 2 Calculated peak areas including noise for each of the seven photopeaks in the synthetic example

Peak energy (keV)	Counts	Relative uncertainty (%)
163.3	3.23×10^4	0.6
276.5	4.00×10^4	0.5
383.0	1.16×10^4	0.9
412.7	1.81×10^3	2.3
584.4	3.82×10^3	1.6
649.0	4.99×10^3	1.4
728.4	9.15×10^3	1.0

Uncertainties were taken as the square root of the value, i.e., Poisson statistics, and are displayed as the relative uncertainty

radionuclides from a gamma-ray spectrum. A multi-gamma standard point source typically used for efficiency calibrations was procured from Eckert and Ziegler Analytics and counted with a high-purity germanium (HPGe) detector. The source was centered approximately 9 cm from the face of the detector. The sample included eleven long-lived radionuclides at known activities: ^{210}Pb (46.5 keV), ^{241}Am (59.5 keV), ^{109}Cd (88.0 keV), ^{57}Co (122.1, 136.5 keV), ^{139}Ce (165.9 keV), ^{203}Hg (279.2 keV), ^{113}Sn (391.7 keV), ^{85}Sr (514.0 keV), ^{137}Cs (661.7 keV), ^{88}Y (898, 1836.1 keV), and ^{60}Co (1173.2, 1332.5 keV). The sample was counted for approximately 20 mins (enough time to get sub-one percent counting statistics), and the spectrum is shown in Fig. 2.

Peaks were located and integrated using Genie 2000 [6]. Nineteen major peaks were found, quantified, and corrected for detector efficiency. Note, not every peak was incorporated into the analysis. Background peaks, e.g., the 1460.8 keV peak from ^{40}K, were excluded based on peak significance (a measure used by Genie to the degree to which a peak stands out from the background and overlapping peaks). Following a peak search[3], the activities at the beginning of counting for each nuclide hypothesized to be present was calculated using Eq. (6).

Results and discussion

This section summarizes the results of the experiments described in Sect. 4.

Synthetic example

Using the counting data in Table 2, Eq. (6) was solved to estimate the activities of each of the seven hypothesized nuclides at the beginning of counting. Table 3 gives the results.

The average relative error in the estimated activities is 2.2 %. The errors are equally distributed between over- and under-estimations, indicating there is no significant bias in the results. When radioactive decay between parent and daughter nuclides is neglected, large errors in the activity of daughter products (nuclides 3 and 5) is observed, and these errors affect other activity estimates because spectral resolutions are incorrectly resolved (see nuclide 4, which interferes with nuclide 3).

---

[3] A full description of our analysis tools is outside the scope of this paper. In short, the peak search relies on comparing peak energies to energies in a database we have generated, populating a response matrix with nuclides that match energies to within calibration uncertainty, and forming the data structures used in Eq. (6).
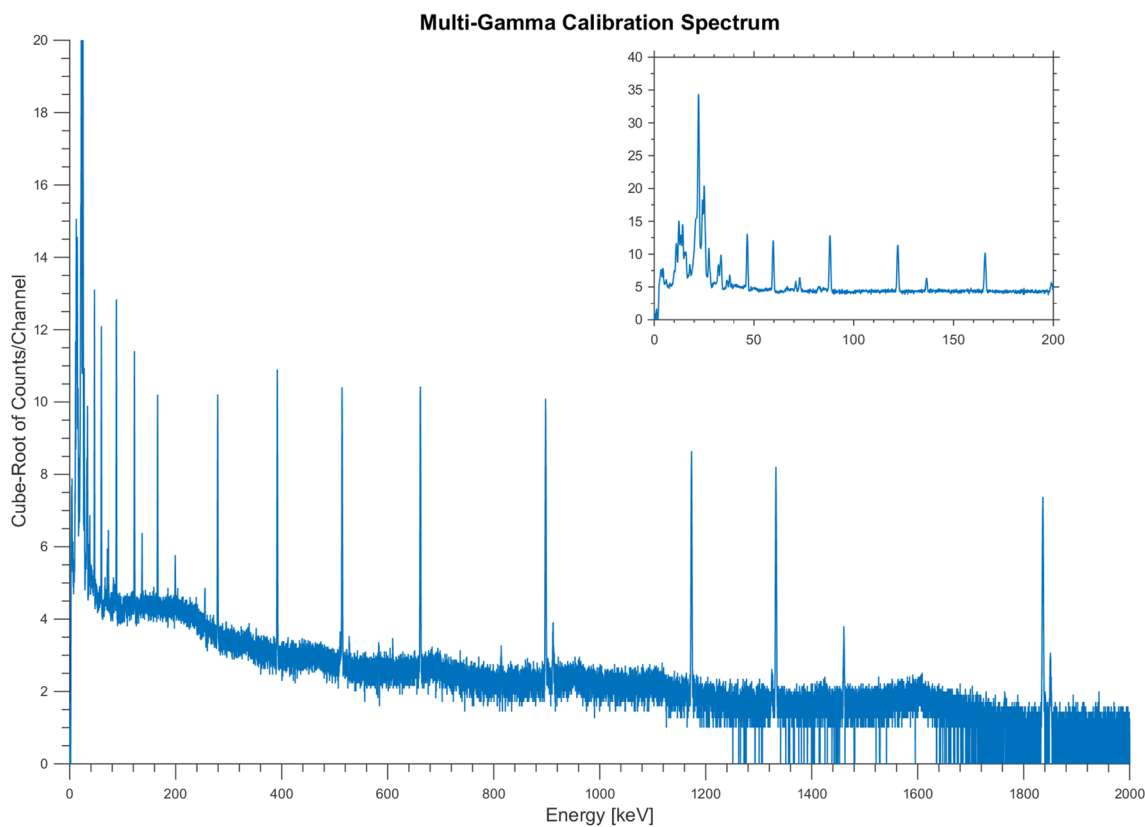
Fig. 2 Spectrum collected of multi-gamma standard using HPGe detector

Table 3 Results from the synthetic experiment described in Sect. 4

Nuclide	Calculated activity (Bq)	Relative error (%)	Estimated activity (no P/D)
1	995.6	−0.44	995.6
2	618.7	+3.11	618.7
3	103.6	+3.57	171.7
4	487.8	−2.45	554.8
5	253.5	+1.38	328.3
6	586.8	−2.20	586.8
7	1.1×10^{-11}	n/a	4.4×10^{-11}

The average relative error is 2.2 % and there is approximately as much over- as under-estimation, indicating there is no significant systematic bias to the results. The last column gives the results when parent/daughter effects are neglected, and signicant errors are observed

The results of this synthetic example suggest that our method for spectrum analysis can quickly resolve multiple spectral interferences as well as correctly quantify decay chains of up to two radionuclides with short half lives relative to the counting time. The method should generalize to longer decay chains; however, this has not been explicitly tested to date.

Long-lived multi-gamma standard

The peak search was performed using a tolerance of 0.15 keV over energies between 25 and 2500 keV. Only considering nuclides with a half-life greater than 3 min, 44 nuclides were hypothesized to be in the sample. Using Eq. (6), the activity of each nuclide at the beginning of counting was calculated. Table 4 gives the results.

All nuclides except ^{241}Am were correctly identified, and the majority of the estimated activities agree well with the known values. Our analysis predicted the presence of ^{237}Pu as opposed to ^{241}Am, as the second, low-intensity photo-peak at 26.3 keV of ^{241}Am was not correctly identified. This is due to an error in the energy of this peak stemming from extrapolation of the energy calibration and overlapping peaks in the low-energy region of the spectrum, which

Table 4 Results of analyzing multi-gamma standard gamma-ray spectrum

Nuclide	Known activity (Bq)	Calculated activity (Bq)
[210]Pb	1690.9 ± 67.6	2057.9
[241]Am	1159.2 ± 6.4	0
[237]Pu	–	1905.0
[109]Cd	2097.0 ± 83.9	736.6
[109]Pd	–	1399.7
[57]Co	80.0 ± 2.8	70.7
[139]Ce	78.8 ± 3.2	38.3
[139]Ba	–	144.2
[203]Hg	116.3 ± 4.7	120.7
[113]Sn	291.2 ± 11.7	287.1
[85]Sr	257.4 ± 10.3	232.3
[137]Cs	372.9 ± 14.9	407.8
[88]Y	544.1 ± 21.8	514.8
[60]Co	445.2 ± 17.8	451.6

Forty-four nuclides were hypothesized to be present, and fourteen were predicted to actually be present. Known activities are reported with a coverage factor of 2, and uncertainties are not calculated by our method at this time

is highly populated by X-rays. On the other hand, [237]Pu has only one strong gamma ray, which may be freely fit to the data in the spectrum.

Quantification for [109]Cd and [139]Ce is flawed due to the inclusion of [109]Pd and [139]Ba, respectively. Both [109]Cd and [109]Pd decay to excited states of [109]Ag, meaning they have the same characteristic photopeaks. Similarly, [139]Ce and [139]Ba both decay to [139]La. These confusions could be resolved using multiple gamma-ray spectra, which would suggest the different half-lives of the competing radionuclides.

The results of this experiment demonstrates that the analysis methodology is promising for rapid radionuclide identification.

Conclusion

We have developed a mathematical framework for the rapid analysis of gamma-ray spectra relying on a numerical optimization approach with an embedded solution to the Bateman equations. Given a list of photopeak energies and net areas, our analysis identifies gamma-emitting radionuclides and quantifies the activity at the beginning of spectral acquisition taking parent/daughter transmutation, radioactive decay, and spectral interferences into account.

Our analysis method presents several advantages over currently available software. We are able to handle an arbitrarily large data library. The eye-readable database developed for this work includes all gamma-emitting nuclides available in the ENSDF dataset. In addition, our approach explicitly models all hypothesized nuclides as decay chains and differentiates between the quantity of a nuclide initially present in the sample and concentration that grows in due to decay of parent nuclide(s). It is possible that nuclides without appreciable gamma emission rates that are coupled to gamma-emitting parent or daughter nuclides may be quantified by enforcement of the correct time behavior. The decay chains may be arbitrarily large, although we have only tested two-nuclide chains to date.

We have tested our method on a synthetic example that demonstrates our ability to quantify radionuclides with complex spectral interferences. We have also summarized an example using real-world data collected of a multi-gamma standard typically used for efficiency calibration. Our novel method successfully identified ten of the nuclides present in the sample and misidentified one; however, additional errant results were reported by the analysis due to identical gamma-ray signatures of nuclides that decay to the same daughter nuclide.

We will continue to add additional logic to make the analysis more robust to real-world effects. These efforts will likely include analysis of multiple gamma spectra to take advantage of changes over time and corrections for self-attenuation of photons within the sample. We are pursuing additional refinements to improve the accuracy of activity estimates and calculate uncertainties in the determined values.

References

1. Gilmore G (2008) Practical gamma ray spectrometry. John Wiley & Sons, West Sussex
2. Chadwick MB, Herman M et al (2011) Nucl Data Sheets 112:2887
3. CVX Research Inc (2011) CVX: Matlab software for disciplined convex programming, version 2.0, http://cvxr.com/cvx
4. Grant M, Boyd S (2008) Graph implementations for nonsmooth convex problems. In: Blondel V, Boyd S, Kimura H (eds) Recent advances in learning and control. Springer, New York, pp 95–110
5. Sturm J (1999) Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones. Optim Methods Softw 11:625–653
6. Canberra Inc. (2009) Genie 2000, version 3.2.1