



Degenerate Pile-up Correction in Pulse Height Spectra from Gamma-ray Spectrometers

Giulia Marcer¹ · Andrea Dal Molin² · Marica Rebai² · Davide Rigamonti² · Marco Tardocchi²

Accepted: 7 May 2024
© The Author(s) 2024

Abstract

One of the primary obstacles faced by spectrometers operating under high counting rates is pile-up, which occurs when two or more events are detected within a timelapse short enough to result in a superposition of the events waveforms. These can not hence be integrated separately in order to get their amplitudes. Piled-up events are typically identified using pile-up rejection or recovery algorithms. In the latter case, the constituent single waveforms and their amplitudes are also restored. However, there are instances in which the pulses overlap so closely that it is impossible to identify the occurrence of pile-up, resulting in the integration of these pulses into a single spurious event. This phenomenon is known as degenerate pile-up. A method to rectify the incorrect reconstruction of degenerate pile-up was developed, based on a statistical approach, which can be directly applied to the pulse height spectra distributions. The approach was tested on a number of synthetic spectra, with counting rates ranging from 20 kHz up to 1 MHz. The recovered spectra were compared to those purely analysed with a pile-up recovery algorithm, demonstrating an improvement of the reconstructed spectrum of several tens of percent when compared to the true synthetic counterpart.

Keywords Gamma-rays · Spectrometer · Degenerate-pile-up · Recovery

Introduction

When scintillation crystals coupled to light collectors are employed as gamma-ray spectrometers, one of the primary obstacles faced under high counting rates is pile-up [1–5]. More generally, this is an issue to be taken into account with every spectrometer. With such detectors, each event corresponds to a pulse in the output signal. In order to trace it back to the gamma-rays energy, each pulse amplitude is generally computed as the waveform maximum or area: the higher the pulse amplitude, the higher the energy of the incoming particle. When pile-up occurs, though, two or more events are detected within a timelapse short enough to result in a superposition of their pulses, making the determination

of the individual event amplitudes non-trivial. A schematic example of pile-up is shown in Fig. 1).

Traditionally and still in most recently developed algorithms, piled-up events are typically identified and resolved using dedicated Pile-Up Recovery Algorithms (PURA), which restore the constituents waveforms and their amplitudes working directly on the overlapped waveforms [1–5]. However, there are instances in which the pulses overlap so closely that it is impossible to even identify the occurrence of pile-up, resulting in the integration of the pulses into a single spurious event. This phenomenon is known as degenerate pile-up. The consequence of degenerate pile-up – as for pile-up in general – is a general loss of events in the low-energy-side of the pulse-height-spectrum (PHS) and a surplus of spurious events in the high-energy one.

A Degenerate pile-Up Correction Algorithm (DUCA) was developed to rectify the incorrect reconstruction of degenerate pile-up events. The algorithm is based on a statistical approach and can be directly applied to the PHS, based only on its acquisition time, without any direct connection with the piled-up waveforms. The algorithm description will follow in section "Correction Algorithm", divided into basic assumptions in section "Basic Assumptions", the

✉ Giulia Marcer
g.marcer@campus.unimib.it

¹ University of Milano-Bicocca, Piazza dell'Ateneo Nuovo, 1, Milan 20126, Italy

² Institute for Plasma Science and Technology, National Council of Research, Via Roberto Cozzi, 53, Milan 20125, Italy

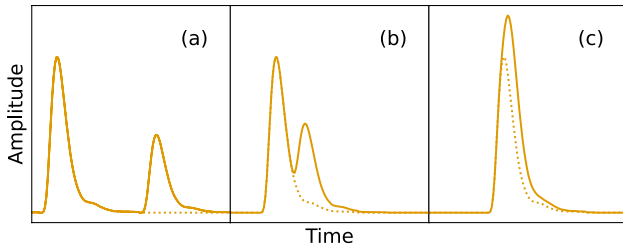


Fig. 1 Schematic example of consecutive events in a spectrometer output signal: **a** no pile-up occurs; **b** the two waveforms overlap but they are still clearly distinguishable; **c** degenerate pile-up: the two events are so close that they appear as a single larger event

actual spectrum correction in section "[Spectrum Correction](#)". Tests of the algorithm with synthetic data follow in section "[Algorithm Tests](#)". A discussions on the algorithm stability is then presented in section "[Algorithm Stability](#)", before getting to the conclusions in section "[Conclusion](#)".

Correction Algorithm

Basic Assumptions

DUCA relies on the assumption that the process of events detection is random enough to follow a Poisson distribution. This assumption allows to compute the probability of detecting two events closer than a certain time τ based only on a single parameter, which is the average events rate R , according to Eq. 1 [6]. DUCA could hence also be extended to other time distributions, prior this probability is known or can be computed based on some known parameters.

$$P_{pu} = e^{-R \cdot \tau} (1 - e^{-R \cdot \tau}) \tag{1}$$

If, instead of a generic time window, τ is the effective time below which the employed PURA is unable to recover events, then P_{pu} is the probability of having degenerate pile-up. The threshold time τ can be determined through the application of DUCA itself. For the moment, τ is assumed to be a known value non-depending on the overlapped pulses energies. This assertion will be justified in section "[Algorithm Stability](#)".

Precisely because of pile-up, though, the measured rate R_m is lower than the true one R . Luckily, it is possible to compute it by solving a transcendental equation:

$$R = - \frac{W[-\tau(R_m + (N_r/T))]}{\tau} \tag{2}$$

where W is the Lambert W function [7], T is the acquisition time and N_r is the number of events which could have eventually been rejected by PURA.

DUCA relies on an other further assumption: it assumes that each time two events overlap, they result in the spectrum as a single event having as amplitude the sum on their amplitudes. This entails that DUCA works better if the PURA integrates the pulses areas for getting their amplitudes instead of using their maxima or performing fits on the pulse shape. The results here discussed were obtained with a PURA [5] using a sort of a middle-way between pulses areas and fits on the waveforms.

Spectrum Correction

Consider a PHS obtained as a PURA output, with measured bins contents $\{m_i\}$. Once normalised to unity, the new bins contents $\{h_i\}$ can be interpreted as a probability distribution function. DUCA works on it bin-wisely. For the i^{th} bin, it computes the probability of getting spurious events P_i^{in} and that of losing true events P_i^{out} . Then it corrects consequently:

$$h_i^* = h_i + P_i^{out} - P_i^{in} \tag{3}$$

where h_i^* is the correct bin probability.

P_i^{out} is the probability of simultaneously getting an event in the i^{th} bin and having it piled-up with at least an other event, which means:

$$P_i^{out} = h_i^* P_{pu} \tag{4}$$

The first bin can not gain spurious events and hence $P_1^{in} = 0$. This allows to correct it as:

$$h_1^* = \frac{h_1}{1 - P_{pu}} \tag{5}$$

For every other bin, on the other hand, computing P_i^{in} requires some math. In addition, it also requires to put an arbitrary threshold on the number of possible piled-up events taken into account. Suppose, for example, that only single pile-up is considered. Then a sum has to be performed over all possible couples of events whose areas sum is equivalent to the current bin and require them to be piled-up. Note that swapping the order of two piled-up events corresponds to consider a different case of pile-up and that both have to be taken into account in the sum.

$$P_i^{in} = \sum_{j < i} h_j^* h_{i-j}^* P_{pu} \tag{6}$$

A note should be pointed out about the finite width of the bins. This point can be better explained with an example. Consider the case in which $i = 7$ and $j = 3$ in Eq. 6. In a not-rebinned spectrum, the j -th bin edges would be 2 and 3 and the $(i - j)$ -th bin edges would be 3 and 4. This means that the sum of two of these events could range from 5 to 7, hence covering the i -th bin itself but also the $(i - 1)$ -th one.

Similarly, when considering $i = 6$ in the previous step, the probability would be shared with the $i = 5$ bin. In DUCA, the bins width is considered narrow enough to neglect this effect.

If, for example, double and triple pile-up are also taken into account, the formula gets recursively more complicated, since the second and third event can in turn be the result of degenerate pile-up. Equation 7 shows how Eq. 6 changes in case both double and triple pile-up are taken into account.

$$P_i^{in} = \sum_{j < i} h_j^* \left[h_{i-j}^* (P_{pu} - P_{pu}^2 - P_{pu}^3) + \sum_{k < i-j} h_k^* (h_{i-j-k}^* (P_{pu}^2 - P_{pu}^3)) + \sum_{l < i-j-k} h_l^* h_{i-j-k-l}^* P_{pu}^3 \right] \tag{7}$$

Note that P_i^{in} always depends only on the bins corresponding to lower amplitudes with respect to the i^{th} itself. Therefore, the correction of the spectrum starts from the 1st bin and moves toward higher amplitudes, one bin at a time.

Once P_i^{in} is known, the i^{th} bin can be corrected according to Eq. 8.

$$h_i^* = \frac{h_i - P_i^{in}}{1 - P_{pu}} \tag{8}$$

Lastly, it is sufficient to multiply $\{h_i^*\}$ for the acquisition time and the acquisition rate R to get the correct bins content $\{m_i^*\}$.

If N is the number of piled-up events, DUCA has a $O(n^N)$ time complexity and hence it could be convenient to perform a moderate rebinning of the spectrum before its application. So, a trade-off is to be considered between running time and accuracy of the reconstructed spectrum.

Algorithm Tests

Synthetic Gaussian Spectrum

A synthetic spectrum with a Gaussian distribution was used to check DUCA behaviour on bins with decreasing content within different orders of magnitude and find its working limits.

One billion waveforms 150 ns long were generated, with 5 ns sampling frequency (30 points each) and double floating points amplitude precision, according to a Poisson distribution in time with $R = 1$ MHz average rate. A Gaussian centered in zero and with 3000 channels as standard deviation was employed as probability distribution.

For the used PURA [5], the effective threshold time τ was 35 ns (see section "Algorithm Stability"), entailing a degenerate pile-up probability of $p_{pu} = 3.3\%$, according to Eq. 1. The output spectrum was then rebinned by a factor 32 before DUCA application.

No source of noise was added in order to not introduce further sources of discrepancies between the true spectrum and the reconstructed one, so that any difference would have only been due to a mis-reconstruction of the spectrum by PURA and DUCA.

In Fig. 2, the true synthetic spectrum is shown together with PURA and DUCA outputs. DUCA was run both considering up to 2 (level 1) or 3 (level 2) simultaneously piled-up events. As expected, the PURA underestimates the number of events in the low-energy-side of the spectrum and finds a surplus in the high-energy one. DUCA, on the other hand, is capable of restoring the original shape if run at the second level. The third level (4 piled-up events) leads to a negligible difference and was hence not even shown in the plot.

Some events are missing in the low energy range of the spectrum because they were below the PURA trigger and not because of not recovered pile-up.

The level 2 DUCA histogram is close to the original one up to around channel 6000, so it can be concluded that DUCA is capable of restoring the correct bins content down to $2 \cdot 10^{-5}$ order of magnitude compared to the most populated bin in the spectrum.

Synthetic Uniform Spectra and Varying Count Rates

In order to quantify the quality of DUCA corrected spectra as a function of the events rate, nine synthetic, uniform spectra one billion events each were generated, each

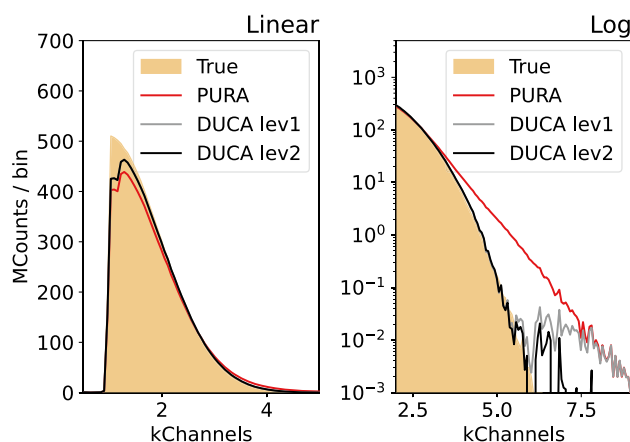


Fig. 2 Synthetic Gaussian spectrum. Comparison between the true one (filled area) and those obtained as PURA (red line) and DUCA outputs (gray and black lines). Linear scale on the left and \log_{10} scale on the right. See the text for description

corresponding to a different rate R , ranging from 20 kHz to 1 MHz. The maximum generated event amplitude was 500. The waveforms characteristics were the same of the Gaussian spectrum. This way, the degenerate pile-up probability ranged between 0.07 to 3.3%. Even this time, the pulses were produced without adding electrical noise. In addition, DUCA was run at its first level (2 piled-up events at most) to fasten the analysis.

For each spectrum, DUCA was applied after PURA and a 32-bin rebinning of the histogram, then both algorithms outputs were compared to the true synthetic spectrum. For the comparison, a simplified version of the Earth Mover's Distance (EMD) was employed [8]. This figure of merit is a measure of the dissimilarity between two histogram distributions, which in this case was computed simply as the sum of the absolute differences between the bins contents of the two histograms, normalized to the original total number of events:

$$\text{EMD} = \frac{1}{N} \sum_i |h_{i1} - h_{i2}| \quad (9)$$

where N is the original total number of events and $\{h_i\}_1$ and $\{h_i\}_2$ are the two histograms under consideration. In short, its value corresponds to the fraction of events out of place in a distribution with respect to a reference one. Since pile-up is a concern when the high-energy-side of the PHS is under study, the out of place events were only considered if in the high-energy-side itself, after the last non-empty bin of the original uniform distribution.

Figure 3 shows the percentage EMD both for PURA and DUCA in the plot above and their percentage ratio in the plot below, as functions of the events rate R . In particular, the metric in Eq. 10 was used for the EMDs comparison.

$$\text{EMDs ratio} = \frac{\text{EMD}_{\text{DUCA}} - \text{EMD}_{\text{PURA}}}{\text{EMD}_{\text{PURA}}} \quad (10)$$

The fraction of out of place events for PURA increases with the events rate up to 1.5% at 1 MHz, while DUCA EMD also increases but its absolute value with respect to PURA EMD saturates down to 80%. It could be useful to point out that the PURA fraction of out-of-place events is lower than the expected fraction of piled-up because only the high-energy-side of the spectrum is taken into account.

Algorithm Stability

As already mentioned in section "Basic Assumptions", the effective threshold time τ of a certain PURA can be estimated with DUCA itself. In principle, because of noise, τ could depend on the amplitudes of the pulses at stake: the higher the amplitudes of the pulses, the higher the

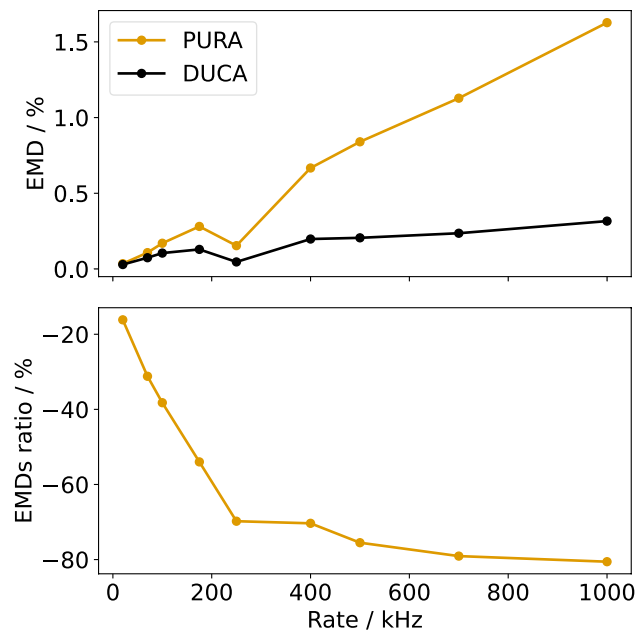


Fig. 3 Percentage Earth Mover's Distances for PURA and DUCA on uniform spectra with varying events rate in the plot above. DUCA percentage improvement with respect to PURA in the plot below

probability for PURA to success. Though, it is still possible to identify an effective τ which minimizes the DUCA output EMD with respect to the original spectrum. This stresses out the importance of simulating a measurement before applying DUCA, if the threshold time is not known a priori.

The first bins of a measured PHS are always empty, being a trigger employed to assess the presence of an event in the signal oscillogram. At most, they may contain events as piled-up events recovered by the PURA@. Though, this is not a concern at all for DUCA. This was proven with a 1 MHz uniformly distributed synthetic spectrum, by setting the first 50 bins to zero. Whether these events were present or not in the spectrum did not affect the high-energy-side if not for less than 1% of the events removed by DUCA.

Regarding the maximum working rate, DUCA limitations are due to its intrinsic probability nature. Whenever the probability of multiple pile-up is very high (at least more than some MHz), also the total number of events would have to be high enough to guarantee the reliability of DUCA application. By the way, DUCA limitations do not represent a concern, since the main restriction is linked to the former PURA application. Whenever the events rate is so high that the probability of getting more than three piled-up events is not negligible, any PURA capability of restoring non-degenerate piled-up events would not be reliable anymore.

Lastly, the algorithm stability with respect to the possible misplacement of some events in the spectrum was tested by changing the seed of the synthetic spectrum random number

generator. This led to a $\pm 4\%$ variation in the EMDs ratio at 1 MHz, thus proving the robustness of DUCA.

Conclusions

Pile-up represents a concern for spectrometers operating at high counting rates. In case of degenerate pile-up, current Pile-Up Recovery Algorithms (PURA) can not identify the occurrence of pile-up.

A Degenerate pile-Up Correction Algorithm (DUCA) was developed to correct PHS for degenerate pile-up. DUCA corrects the spectra bin-wisely and relies on a single parameter, which is the PURA pile-up discrimination threshold time. In case it is not known, this threshold time can be estimated with DUCA itself with the aid of synthetic spectra.

DUCA is based on two assumptions: the detected events follow a Poisson distribution in time; in case of degenerate pile-up, the spurious event amplitude corresponds to the sum of the piled-up events amplitudes. The algorithm does not intrinsically present concerning rate working limits, nor is affected by the presence of a trigger at low energies.

It was tested on synthetic uniform spectra ranging from 20 kHz up to 1 MHz, always improving the Earth Mover's Distance with respect to the original one down to -80% at 1 MHz.

DUCA can hence be employed to correct high-counting-rates spectra from degenerate pile-up at least up to 1 MHz.

Author Contributions A.D.M. pointed out the existence of degenerate pile-up while the research group was working on a certain project. G.M. developed the algorithm for the degenerate pile-up correction. A.D.M. and M.T. helped to develop the algorithm and test it on synthetic data. M.R. and D.R. helped interpreting the results of the analysis on the synthetic data. G.M. wrote the main manuscript text. A.D.M., D.R., M.R. and M.T. reviewed the manuscript.

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement. This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 - EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Availability of data Both the implemented algorithm and all synthetic data shown in this work can be obtained by asking to the corresponding author via e-mail.

Declarations

Conflict of interest The authors declare no Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. M.E. Hammad et al., Pile-up correction algorithm for high count rate gamma ray spectroscopy. *Appl. Radiat. Isot.* **151**, 196–206 (2019)
2. T. Petrovic et al., Efficient reduction of piled-up events in gamma-ray spectrometry at high count rates. *IEEE Trans. Nucl. Sci.* **61**, 584–589 (2014)
3. M. Bolic et al., Pileup correction algorithms for very-high-count-rate gamma-ray spectrometry with NaI(Tl) detectors. *IEEE Trans. Instrum. Meas.* **59**, 122–130 (2009)
4. J. Zuo et al., High count rate gamma ray pile-up correction method of NaI(Tl) detector based on deconvolution method. *Nucl. Instrum. Methods Phys. Res. A* **1061**, 169135 (2024)
5. G. Marcer et al., A new dedicated signal processing system for gamma-ray spectrometers in high power deuterium-tritium plasma scenarios in tokamaks. *Rev. Sci. Instruments* **93**, (2022)
6. G.F. Knoll, *Radiation detection and measurement* (Wiley, 2010)
7. E.W. Weisstein, Lambert w-function (2002). <https://mathworld.wolfram.com>
8. Y. Rubner et al., The earth mover's distance as a metric for image retrieval. *Int. J. Comput. Vis.* **40**, 99–121 (2000)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.