Restoration of Poissonian events that were missed due to extending dead time

V. Horvat and J.C. Hardy

If a known extending dead time is imposed to a Poissonian sequence of events, some of the events will be lost. We are trying to develop a method of replacing those lost events by *simulated* events in a way that guarantees (within precision of the best available measurements) that the resulting event sequence is purely Poissonian and statistically indistinguishable from the original sequence of events. This problem proved not to be as trivial as it may seem, even in the simplest case (considered here), in which the ideal rate of the original Poissonian events is constant and known.

In this report we describe a solution to this problem, along with some potentially interesting properties of the event distributions that were discovered along the way. The method of event restoration shown here is found to be effective and appropriate, and it can be set to work within any desired level of accuracy. For brevity, we will use the term *event restoration* instead of the more meaningful term *replacement of events (lost due to dead time) by a statistically equivalent set of simulated events*.

The method of event restoration proposed here is explained based on an example of a data set with approximately 60 million events simulated assuming constant ideal event rate $\rho = 10\ s^{-1}$, and partitioned into one thousand samples, each having 100 channels spanning the duration of 100 s. An extendable dead time per event, τ_e, is then imposed to this event sequence. Several values of τ_e up to as high as 120 ms, were used in order to prove the applicability of the method even in the extreme situations. Specifically, at τ_e = 120 ms, the expected fraction of surviving events [given by $\exp(-\rho\tau_e)$] is only about 30%.

Given that the sequence of known live times associated with the surviving events is Poissonian, it may seem appropriate to map all consecutive dead-time intervals into a single continuous time interval, fill that single interval with simulated events at the nominal ideal event rate, and then map each dead-time interval together with its events back into the interval's original place in the time sequence. Even though this process produces a statistically correct number of restored events, the resulting time sequence is non-Poissonian and can be identified as such by looking at the distribution of its time intervals or by looking at the variance of the resulting number of events per channel of a chosen size as a function of time. Furthermore, reapplying the dead time to this sequence of events yields a secondary event set significantly different from that obtained originally. Nevertheless, a method similar in concept to this one, referred to as the "shadow method", has been used in the analysis of muon-decay spectra [1].

In order to restore the lost events so that the resulting event time sequence is Poissonian, it is necessary to observe and respect the following four properties of the secondary-event sequence:

(*i*) a dead-time interval cannot be shorter than τ_e,
(*ii*) if a dead-time interval length is τ_e, then no events have been lost,
(*iii*) if a dead-time interval is longer than τ_e, then one lost event has originally occurred at time τ_e before the end of the dead-time interval, and eventually,

(*iv*) any number of events may have occurred before that, provided that the differences between arrival times of the consecutive lost events are less than τ_e.

While incorporating properties (*i*) through (*iii*) is trivial, simulating the events that may have occurred before the last lost event in a dead-time interval longer than τ_e proved to be challenging. Specifically, if one event is generated at a time, then the arrival time of the last lost event (that is known to have happened) relative to the previous event (obtained in a simulation) will not be properly distributed. In fact, longer intervals will occur more often than shorter intervals, which is opposite to what Poissonian distribution requires.

Furthermore, simulating a series of events restricted to time intervals shorter than τ_e results in a wrong time-interval distribution and a significantly overestimated number of restored events. Although the cumulative probability of a truncated probability density scales with that of an unrestricted probability density for a single time interval, the two do not scale when a sequence involving more than one time interval is involved. Specifically, density of the probability of having exactly one lost event in a dead-time interval, t_d (which occurs at t_d - τ_e), is given as a function of t_d by

$$dP_{1L}/dt_d = \begin{cases} 0 & 0 < t_d \leq \tau_e \\ \rho \exp[-\rho(t_d-\tau_e)] / [1-\exp(-\rho\tau_e)] & \tau_e < t_d \leq 2\tau_e \\ 0 & t_d > 2\tau_e \end{cases} \qquad (1)$$

which scales with $\rho \exp[-\rho(t_d-\tau_e)]$. However, the density of the probability of having exactly two lost events in a dead-time interval, t_d (one of which occurs at t_d - τ_e), is given by

$$dP_{2L}/dt_d = \begin{cases} 0 & 0 < t_d \leq \tau_e \\ \rho^2(t_d-\tau_e) \exp[-\rho(t_d-\tau_e)] / [1-\exp(-\rho\tau_e)]^2 & \tau_e < t_d \leq 2\tau_e \\ \rho^2(3\tau_e-t_d) \exp[-\rho(t_d-\tau_e)] / [1-\exp(-\rho\tau_e)]^2 & 2\tau_e < t_d \leq 3\tau_e \\ 0 & t_d > 3\tau_e\ , \end{cases} \qquad (2)$$

which does scale with $\rho^2(t_d-\tau_e) \exp[-\rho(t_d-\tau_e)]$ for $\tau_e \leq t_d \leq 2\tau_e$ because in this case the first event can occur anywhere in the range between 0 and the second event (which occurs at t_d - τ_e), but this scaling does not extend to $2\tau_e \leq t_d \leq 3\tau_e$ because in this case the occurrence of the first event is restricted to (t_d-$2\tau_e$, τ_e) as a consequence of the fact that time intervals between consecutive events larger than τ_e are ruled out. Similar formulas for multiple event loss could be derived and used, but this approach was found to be impractical because the formulas become more complex as the number of events lost increases, and because there is no limit to the number of events that could have been lost.

To solve this problem, an appropriate and effective approach is suggested here, in which the entire sequences of unrestricted events are simulated repeatedly for each dead-time interval longer than τ_e, until a sequence is found that has all arrival-time differences between consecutive events less than τ_e and also features an event occurring, within the selected tolerance, at time τ_e before the end of the dead-time interval. This method was found capable of producing an event set statistically indistinguishable from the original event set and therefore, in the resulting histogram of the number of events per channel of a chosen size as a function of time, the best estimate of the variance of the number of events in each channel equals the number of events in that channel. This was verified to be true. It was also verified that the distribution of time intervals between consecutive events is exponential and corresponding to the nominal ideal rate, as expected.

However, there are several caveats associated with the way the tolerance is defined in this method. Although a tighter tolerance will lead to a more accurate result, it will also increase the computation time. Also, a tolerance imposed in a biased way produces a biased result and requires tighter restrictions (leading to more computation time) in order to yield a result with the desired accuracy. Therefore, the goal is to define tolerance in an unbiased way, in order to obtain the desired accuracy faster, *i.e.*, by imposing restrictions that are less tight. To do that properly, one must be aware of the discontinuities in the distribution functions for the dead-time interval durations and the associated numbers of events, as well as the fact that the latter are combinations of compactly supported functions

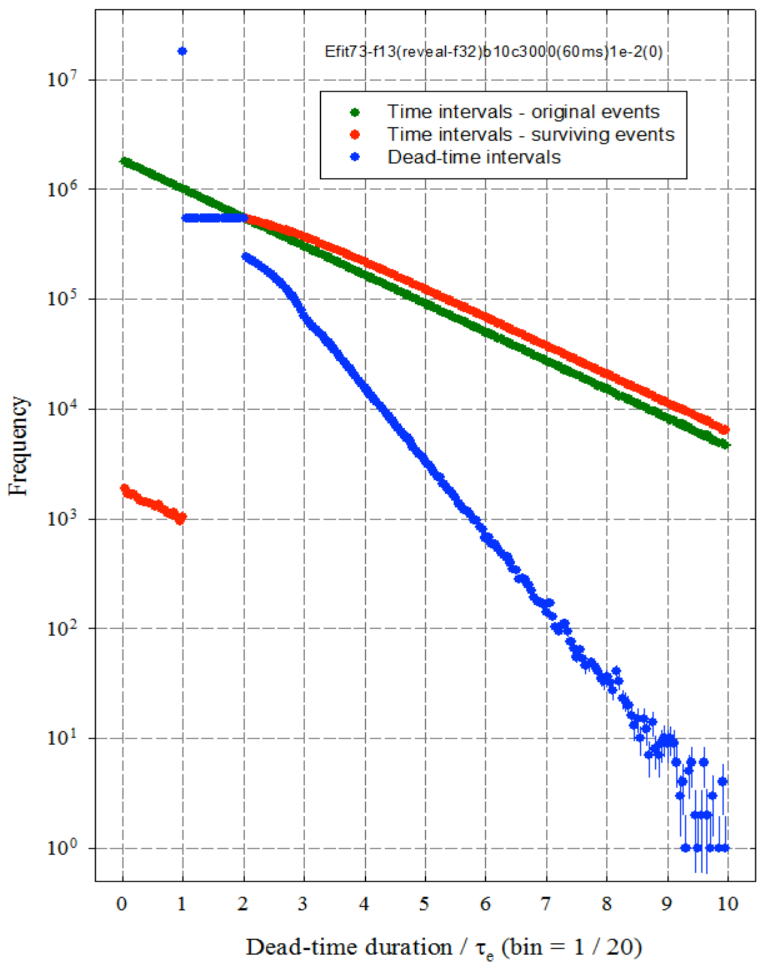

FIG. 1. Distribution of dead-time interval durations in units of τ_e, shown with a bin size of 1/20 (blue circles), and compared with the corresponding distribution of arrival-time-differences between consecutive events that survived the imposition of extendable dead time (red circles) and those that were present before the imposition of dead time (green circles). In this example

that are non-zero only over a domain extending between two characteristic integer multiples of τ_e and that their derivatives have discontinuities at the integer multiples of τ_e within their domain. [Examples are the distributions defined by Eqs.(1-2)].

The graphs in Figs. 1-2 are given in order to illustrate these facts. Fig. 1 shows the distribution of dead-time interval durations in units of τ_e, with a bin size of 1/20. It demonstrates that the distribution of dead-time intervals (shown in blue) is flat for $\tau_e < t_d \leq 2\tau_e$. For $0 < t_d \leq \tau_e$, the distribution is a delta function, $\delta(t_d-\tau_e)$, since, by definition, there can be no dead-time intervals shorter than τ_e. For $t_d > 2\tau_e$ the distribution seems continuous, but it obviously changes its dependence on t_d/τ_e at $t_d = 3\tau_e$. The same is expected to occur at $t_d/\tau_e = 4, 5, 6, \ldots$, but it is less obvious. The remaining two distributions shown in Fig. 1 are well known and understood [2], noting that the red circles between 0 and 1 correspond to the first events in the samples, which are not affected by the dead time, and therefore reflect the distribution shown with the green circles. Also note that the blue circles in the flat part of the distribution overshadow

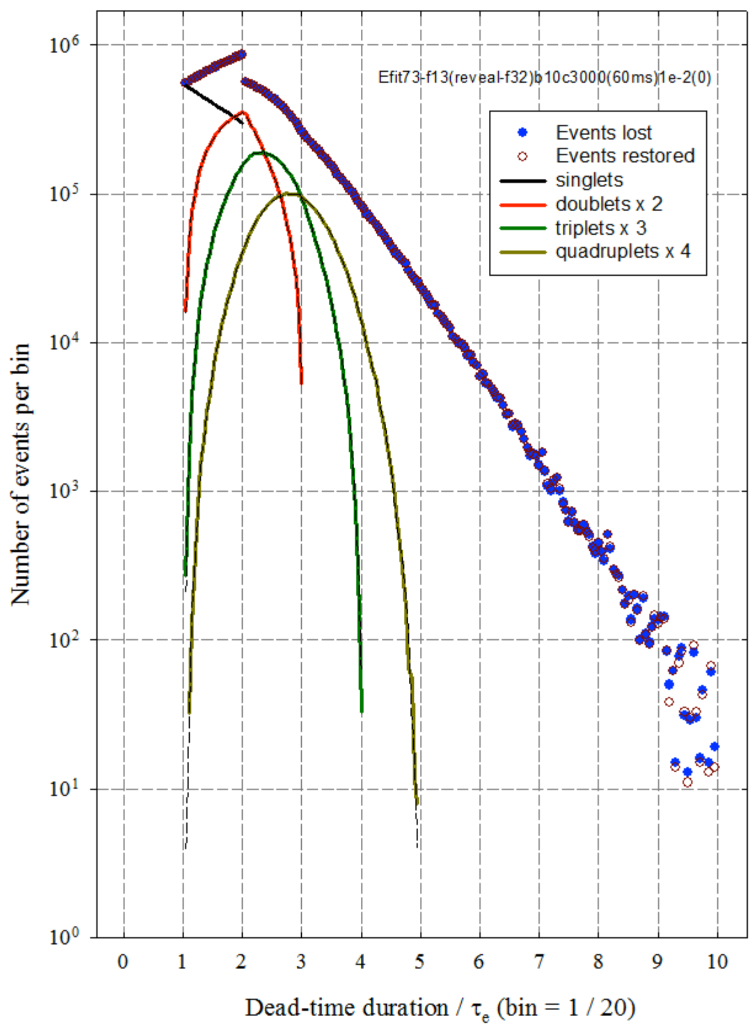

FIG. 2. Distribution of the number of events lost (or restored using the method proposed here) as a function of dead-time-interval duration in units of τ_e, shown with a bin size of 1/20. The solid blue circles and hollow brown circles show the total number of events lost and restored, respectively, while the lines show the contributions based on the number of events lost or restored in the dead-time interval.

the red circles. The overlap is expected to be exact in the central limit, but otherwise, the events contributing to them are not necessarily the same.

Fig. 2 shows the distribution of the number of events lost as a function of dead-time-interval duration in units of τ_e, with a bin size of 1/20. In addition, it shows the breakdown of contributions based on the number of events per dead-time interval for up to and including four events per interval. The sum of all these contributing distributions without their multipliers results in the distribution shown in Figure 1 with the blue solid circles, except, of course, the point at $t_d = \tau_e$, since in that case the dead time interval is present, but no events are lost. The distributions corresponding to the events restored using the method proposed here are the same as those corresponding to the lost events, although the numbers of lost and restored events in a given dead-time interval are not necessarily the same.

It should be pointed out that, while testing this method using a simulated set of events, one must make sure that the program used for the event simulation and the program used for the analysis of the secondary set of events employ different sequences of random numbers (generated using different seeds).

[1] Kevin R. Lynch, Nucl. Phys. B - Proceedings Supplements, **189**, 15 (2009).
[2] J.W. Mueller, Nucl. Instrum. Methods Phys. Res. **A301**, 543 (1991).