

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Karvanen, Juha; Niilo-Rämä, Mikko; Sarén, Jan; Kärkkäinen, Salme

**Title:** Estimating mean lifetime from partially observed events in nuclear physics

**Year:** 2022

**Version:** Published version

**Copyright:** © 2021 the Authors

**Rights:** CC BY 4.0

**Rights url:** <https://creativecommons.org/licenses/by/4.0/>

**Please cite the original version:**

Karvanen, J., Niilo-Rämä, M., Sarén, J., & Kärkkäinen, S. (2022). Estimating mean lifetime from partially observed events in nuclear physics. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 71(1), 3-26. <https://doi.org/10.1111/rssc.12519>

# Estimating mean lifetime from partially observed events in nuclear physics

Juha Karvanen<sup>1</sup>  | Mikko Niilo-Rämä<sup>2</sup> | Jan Sarén<sup>3</sup> |  
Salme Kärkkäinen<sup>1</sup>

<sup>1</sup>Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, Finland

<sup>2</sup>Finnish Institute for Educational Research, University of Jyväskylä, Jyväskylä, Finland

<sup>3</sup>Department of Physics, University of Jyväskylä, Jyväskylä, Finland

## Correspondence

Juha Karvanen, Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, Finland.  
Email: juha.t.karvanen@jyu.fi

## Abstract

The mean lifetime is an important characteristic of particles to be identified in nuclear physics. State-of-the-art particle detectors can identify the arrivals of single radioactive nuclei as well as their subsequent radioactive decays (departures). Challenges arise when the arrivals and departures are unmatched and the departures are only partially observed. An inefficient solution is to run experiments where the arrival rate is set very low to allow for the matching of arrivals and departures. We propose an estimation method that works for a wide range of arrival rates. The method combines an initial estimator and a numerical bias correction technique. Simulations and examples based on data on the alpha decays of Lutetium isotope 155 demonstrate that the method produces unbiased estimates regardless of the arrival rate. As a practical benefit, the estimation method enables the use of all data collected in the particle detector, which will lead to more accurate estimates and, in some cases, to shorter experiments.

## KEYWORDS

design of experiments, missing data, noisy binary search, Poisson process, radioactive decay

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* published by John Wiley & Sons Ltd on behalf of Royal Statistical Society

# 1 | INTRODUCTION

Radioactive decay is the textbook example of a Poisson process in time (Cox & Lewis, 1966). The statistical model can be derived directly from the laws of physics: the decay events are independent given the decay rate and the lifetimes follow an exponential distribution with a mean equal to the inverse of the decay rate. The mean lifetime is an important characteristic of particles to be identified in nuclear physics since it is sensitive to the structure of underlying quantum mechanical states. For example, a large change in nuclear structure from the initial to the final state causes a radioactive alpha decay longer than predicted by a simple model (Geiger & Nuttall, 1911). The estimation of the mean lifetime under different experimental setups is a recognized statistical problem in particle physics (Lyons, 2008).

In many modern experiments, radioactive species are produced continuously and decays are measured simultaneously. This is enabled by state-of-the-art particle detectors and data-acquisition systems which can identify the arrivals of single radioactive nuclei to a detector as well as their subsequent radioactive decays (Lazarus et al., 2001; Page et al., 2003). The arrivals and the decays, called departures for generality, are unmatched meaning that there is no physical way to link an arrival and a departure to each other. Only if the arrival rate is low and the mean lifetime is short, we are able to say with a high probability that a certain arrival and departure form a pair because only one departure has been observed between two consecutive arrivals.

Challenges arise when some arrivals or departures are mislabelled or partially observed. In this paper, we focus on experiments where it is not possible to catch all departures due to the structure of the particle detector. For example, the alpha radioactive nuclei created in fusion reactions and implanted to a detector are typically very close to the surface of the detector which implies that about half of departures are not observed due to the escape of the alpha particle from the detector. Combined with the lack of matching, we are left with an unequal number of arrival and departure times that cannot in general be linked to each other. Estimating the mean lifetime from this kind of data is the problem addressed in this paper.

Experimenters have commonly preferred setups where the arrival rate is low compared with the decay rate because then matching an arrival to a departure before the next arrival causes only a negligible bias (e.g. Herzán et al., 2015, figure 12). However, this approach may lead to inefficient use of the measurement capacity. The matching is not possible for a high arrival rate but on the other hand the number of events per second is high. The motivation of this paper is to propose a method, which can also be used with high arrival rate in the case of unmatched arrivals and departures.

First we show that the matching of arrivals and departures is not necessary for the estimation of the mean lifetime. Next, we present a simple estimator for the case where all arrivals and departures are observed but not matched. When the departures are only partially observed, we use the same estimator after the thinning of the arrivals and apply a numeric bias correction technique. Simulation experiments demonstrate that the proposed method produces unbiased estimates regardless of the arrival rate.

The paper is organized as follows. In Section 2 the details of measurement device and data collection procedure are explained. In Section 3 it is shown that the matching of arrivals and departures is not needed in the estimation. Section 4 describes the proposed methodology for partially observed arrivals and departures. In Section 5, the implementation of the proposed estimation method is described. In Section 6, the properties of the proposed estimator are studied in simulations. In Section 7, the method is applied to  $^{155}\text{Lu}$  (Lutetium isotope 155) alpha decay data. The practical implications and future directions are discussed in Section 8.

## 2 | DATA

### 2.1 | Technical description of the physical data collection

The data were collected in the Accelerator laboratory of University of Jyväskylä in an experiment aiming to measure charged particle radioactivity of  $^{160}\text{Os}$  produced in a fusion evaporation reaction (unpublished data). A beam of 304 MeV  $^{58}\text{Ni}$  from the K130 cyclotron (Liukkonen, 1993) was used to bombard a self-supporting  $^{106}\text{Cd}$  target of thickness around  $1\text{ mg/cm}^2$ . The average beam intensity of 3 pA ( $19 \times 10^9$  particles/s) was used over 14.6 h. The radioactive  $^{155}\text{Lu}$  was one of the products of the fusion evaporation reaction.

Fusion evaporation residues (recoils, called arrivals in this paper) were separated from the unreacted primary beam with the recoil-mass spectrometer called MARA (Sarén, 2011) employing electric and magnetic fields. At the focal plane of MARA the recoils are detected first with a gaseous multi-wire proportional counter (MWPC) transmission detector and then implanted into a double-sided silicon strip detector (DSSD) consisting of  $192 \times 72$  strips in horizontal and vertical directions respectively. The full size of the DSSD is  $128 \times 48\text{ mm}^2$  (width  $\times$  height). Ionizing particle causes a local simultaneous signal into both horizontal and vertical strips of the DSSD and therefore the event can be assigned to a virtual pixel. The transmission position in horizontal and vertical directions of a recoil flying through the MWPC detector is obtained with time to amplitude converter (TAC) modules utilizing the time differences between charge signals from opposite ends of the wireplanes read through delay lines. The directions of the decay products (departures) of the implanted recoils (arrivals) are random and the DSSD covers half of the possible directions of decays, which leads to the detection rate of 50%. Figure 1 presents a schematic illustration of the experimental setup.

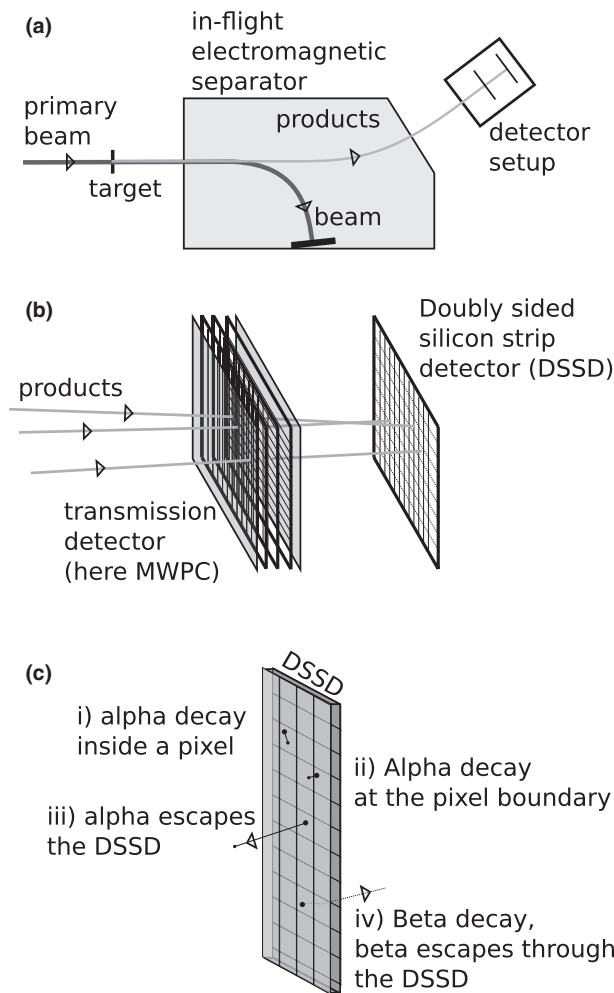
All detector events in MWPC and DSSD are recorded by a self-triggering data-acquisition system employing analogue to digital converter (ADC) modules and field-programmable gate array (FPGA) chips. The kinetic energy of a recoil implanted into the DSSD and the radioactive decay of a recoil implanted earlier are calculated in the FPGA chips running a moving window deconvolution algorithm for the voltage signal generated by charge integrating preamplifiers.

Events observed in the DSSD are classified into three categories: recoils (arrivals), decays (departures) and background. A DSSD event is a recoil if it was also observed in the MWPC detector and the DSSD energy and the time of flight between MWPC and DSSD is inside a two-dimensional gate and both horizontal and vertical positions were recorded successfully at MWPC. Due to the last condition, the probability to classify a recoil as background was 8.7%. A DSSD event without any signal at MWPC and having an energy between 500 keV and 10,000 keV is classified as a decay event which can be in this reaction an alpha, beta or proton radioactivity. All the rest of events are classified as background.

It is possible to continue measuring decays even after the fusion evaporation reaction has ended. This means that right censoring due to the end of the follow-up can be avoided.

### 2.2 | Dataset 1

Dataset 1 is a subset of the full dataset consisting of  $^{155}\text{Lu}$  decay events related to the alpha energy of 7390 keV. The subset was deliberately chosen so that the arrivals and departures can be matched with a high probability. A record containing the time of an  $^{155}\text{Lu}$  alpha decay, the time of the previous recoil in the same DSSD pixel and the horizontal and vertical strip



**FIGURE 1** A schematic view of the production, separation and detection of products of a nuclear reaction when an in-flight recoil separator is used. (a) An almost continuous beam of primary particles from an accelerator hits a target foil. The electromagnetic recoil separator selects the products and transports them to a detector setup. (b) A typical focal plane detector setup consists of a transmission detector (typically multi-wire proportional counter) and a Double-sided silicon strip detector (DSSD). The products are implanted into the surface of the DSSD and the position of the implantation is registered. The arrivals referred in the text form a subset of these events. (c) A radioactive decay of an implanted product can be observed later and can be correlated spatially and temporally to the implantation. Four different decay events are illustrated: (i) a full energy alpha decay takes place in a pixel defined by orthogonal strips of the DSSD, (ii) decay occurs at the pixel boundary and part of the decay energy might be lost, (iii) an alpha particle escaping from the surface of the DSSD and only part of its energy is registered and (iv) the product undergoes a beta decay and the beta escapes through the detector (a beta particle has much higher range in silicon compared to an alpha particle). The departures in the text are of the case (i) events, the cases (ii) and (iii) describe two different mechanism to miss a departure and the case (iv) illustrates a different decay mode creating background events

number was written into a file if the time between these events was less than 1.5 s and there were no other decay events in between. The arrivals misclassified as background were ignored. The dataset contains 16,320 matched arrival–departure pairs and lifetimes obtained as the differences of the departures and the arrivals. Lifetimes shorter than approximately 0.008 ms are

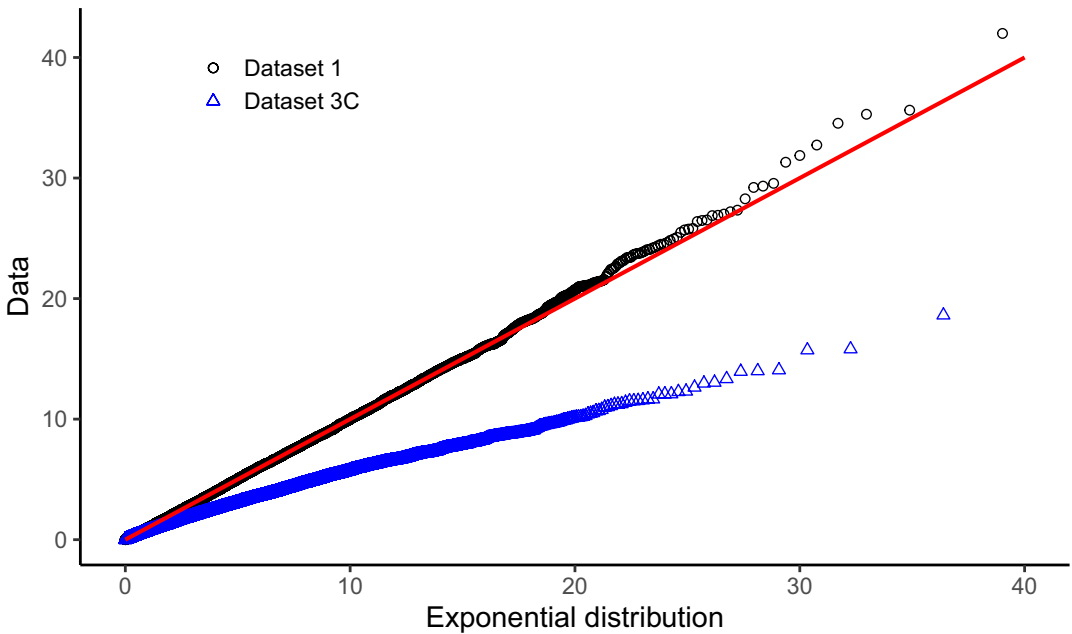
not detectable due to technical limitations, that is, the data are left truncated at 0.008 ms. The time of arrivals varies from 0 ms to 52,577,456 ms (about 14.6 h) implying that the arrival rate is 0.000310 events/ms.

The lifetimes contain some outliers, that is, extremely long lifetimes, which are assumed to be not real. We choose 80 ms to be the cut-point on the basis of a visual inspection and remove 251 lifetimes being longer than the cut-point. Thus, the data are truncated to the interval [0.008 ms, 80 ms]. After the truncation, the number of arrivals is 16,069 and the arrival rate is 0.000306 events/ms. These 16,069 records are called Dataset 1.

The lifetimes are modelled by the exponential distribution. This assumption is supported by the data even after the truncation. The mean of the lifetimes, 3.760 ms, is very close to their standard deviation 3.798 ms and the form of the empirical distribution closely resembles the exponential distribution (Figure 2). For the exponential distribution with the mean lifetime 3.760 ms, the probability mass outside the range [0.008 ms, 80 ms] is only about 0.0021. Thus, it seems justifiable to treat the lifetimes as independent observations from an exponential distribution.

### 2.3 | Datasets 2, 3 and 4

Datasets 2, 3 and 4 are derived from Dataset 1 for illustrative purposes by breaking the matching, modifying the arrival rate and removing a part of the departures. This allows us to compare the estimates obtained under different conditions to those obtained from Dataset 1. The potential misclassification of arrivals was not considered because the focus was in partially observed departures.



**FIGURE 2** The quantile–quantile plot of Dataset 1 and the fitted exponential distribution contrasted with Dataset 3C. The pseudo-lifetimes for Dataset 3C are obtained applying the thinning and matching explained in Section 5.2

Starting with arrival times  $a_1, \dots, a_{16069}$  and the corresponding lifetimes  $x_1, \dots, x_{16069}$  of Dataset 1, we derive modified arrival and departure times as follows

$$\begin{aligned} a'_i &= a_i/c \\ d'_i &= a'_i + x_i, \end{aligned}$$

where  $c$  is a positive scaling constant. Dataset 2 is a densified, unmatched variant of Dataset 1. The arrival times are divided by  $c = 1000$  and the departure times are recalculated by adding the lifetimes to the modified arrival times. When the lifetimes are removed, Dataset 2 contains only arrival and departure time instances without any direct information on the matching or lifetimes.

Datasets 3A–3F are derived from Dataset 1 by first randomly removing each departure with probability 0.5 and then applying densification. These datasets contain 16,069 arrivals and 8042 unmatched departures. The scaling constants  $c = 10$ ,  $c = 100$ ,  $c = 1000$ ,  $c = 10,000$ ,  $c = 100,000$  and  $c = 10^6$  are used to densify Datasets 3A, 3B, 3C, 3D, 3E and 3F respectively. The main goal is to estimate the mean lifetime from Datasets 3A–3F.

Datasets 4A–4F are derived in a similar way as Datasets 3A–3F but each departure is removed with probability 0.75 instead of 0.5. Datasets 4A–4F are used to demonstrate that the proposed estimation method works even if the detection probability is lower than in the real-life experiment. Altogether we have 14 datasets for the analysis.

### 3 | ESTIMATION METHODS FOR COMPLETE DATA

#### 3.1 | Classical setup with matched pairs

The process of interest is defined by two independent Poisson processes in time: an arrival process with arrival rate  $\lambda_a$ , and a lifetime process with decay rate  $\lambda$  and mean lifetime  $\mu = \lambda^{-1}$ . We observe arrival times  $a_1 < \dots < a_n$  and the corresponding departure times  $d_1 < \dots < d_n$ . As the arrivals and departures are matched, we can calculate the lifetimes directly as  $x_i = d_i - a_i$ . The lifetimes  $x_1, \dots, x_n$  are independent observations from an exponential distribution with the mean lifetime  $\mu$ . This is the setup for Dataset 1.

As there are no censored observations, the likelihood function is given by:

$$L(\mu) = \prod_{i=1}^n \frac{1}{\mu} e^{-x_i/\mu}. \quad (1)$$

The maximum likelihood estimator for  $\mu$  is then

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

with the standard error

$$SE(\hat{\mu}) = \frac{\hat{\mu}}{\sqrt{n}}. \quad (3)$$

### 3.2 | Unmatched arrivals and departures

Now, let us assume that all arrival and departure times are observed, but the information on the matching between arrivals and departures is missing. This is the setup for Dataset 2. In practice, the experiment has a certain activity period when all arrivals are observed. After this period, new arrivals do not occur but departures can still be observed. The follow-up period is the activity period extended with a sufficiently long post-activity period to catch all departures with a high probability. The activity time refers to the length of the activity period and the follow-up time to the length of the follow-up period.

The lack of matching may appear as a serious problem at first. Attempts to enumerate all possible ways to match arrivals and departures quickly fail because of combinatorial explosion. Combinatorics can be completely avoided writing estimator (2) in the form

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^n (d_i - a_i) = \frac{1}{n} \left( \sum_{i=1}^n d_i - \sum_{i=1}^n a_i \right), \quad (4)$$

where only the sum of the departure times and the sum of the arrival times are needed to estimate  $\mu$ . We see that the matching does not give any additional information on  $\mu$  and the maximum likelihood estimator is the same as in the classical setup. The standard error of estimator (4) is given by Equation (3).

## 4 | ESTIMATION METHODS FOR INCOMPLETE DATA

In the more complex case present in Datasets 3A–3F, arrivals and departures are unmatched and on average half of the departures are not observed. This situation is illustrated in Figure 3. The observed data consist of  $n$  arrival times  $a_1, \dots, a_n$  and  $m < n$  unmatched departure times  $d_1, \dots, d_m$ .

A straightforward extension of estimator (4) for data with partially observed departures is

$$\frac{1}{n} \left( \sum_{j=1}^m 2d_j - \sum_{i=1}^n a_i \right), \quad (5)$$

where the constant 2 is used to compensate for the unobserved departures. The crucial difference compared with estimator (4) is that a half of the arrival–departure pairs are replaced by randomly chosen arrivals and departures. It follows that the variance of Equation (5) is larger than the variance of Equation (4) because the variance of  $d_j - a_i, j \neq i$  is larger than the variance of  $d_i - a_i$ . In the Appendix we show that the variance of estimator (5) increases as a function of  $n$  and thus the estimator is not usable.

The likelihood inference requires that the arrivals are thinned, that is, some arrivals are removed so that the number of remaining arrivals equals the number of observed departures  $m$ . The thinning is defined by selection indicators  $s_1, \dots, s_n$  for which  $s_i = 0$  if arrival  $a_i$  is removed and  $s_i = 1$  otherwise. Here, thinning  $\mathbf{s} = (s_1, \dots, s_n)$  is called compatible with arrivals  $a_1, \dots, a_n$  and departures  $d_1, \dots, d_m$  if



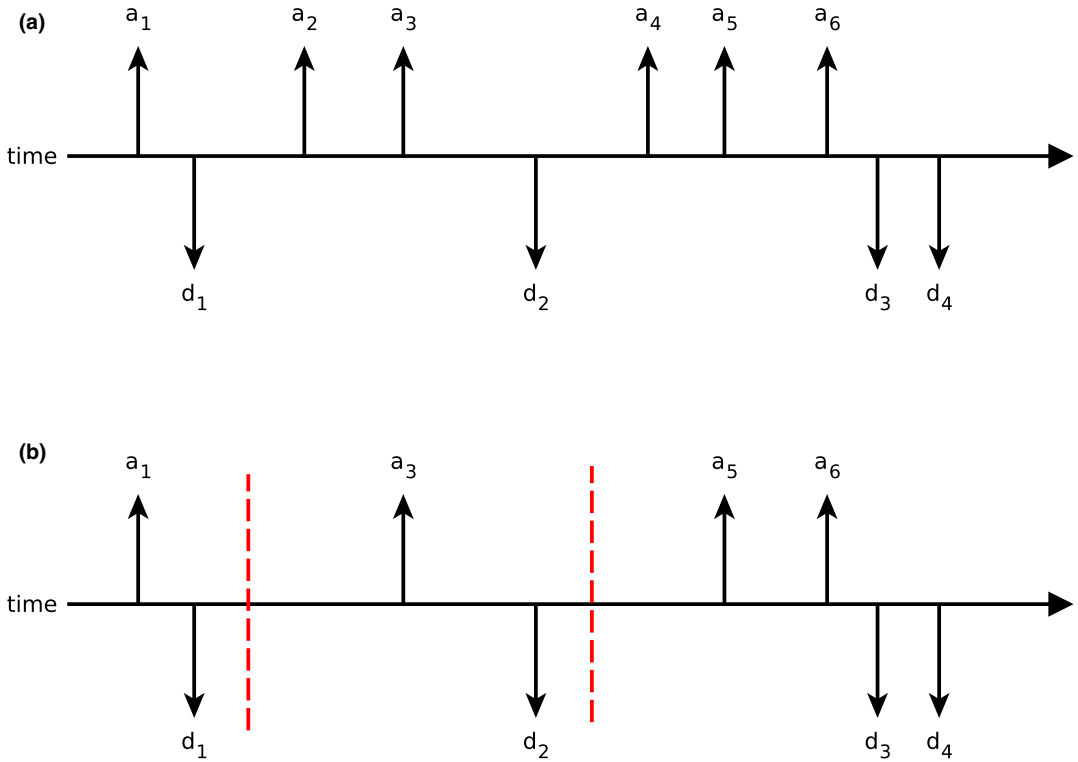


FIGURE 3 An illustration of the process of unmatched arrivals and departures when on average half of the departures are missing. Arrows pointing upward represent arrivals and arrows pointing downward represent departures. The activity period ends somewhere between  $a_6$  and  $d_3$  and the follow-up period ends a long time after the last observed departure  $d_4$ . (a) The original data containing six arrivals and four departures. (b) The data after the division into blocks and the thinning of the arrivals. The vertical dashed lines show the cut-points between the blocks

$$\sum_{i=1}^n s_i = m$$

and for any  $t$  it holds

$$\sum_{i=1}^n s_i I(a_i < t) \geq \sum_{j=1}^m I(d_j < t),$$

that is, after the thinning, the cumulative number of departures never exceeds the cumulative number of arrivals. A compatible thinning guarantees that it is possible to match the arrivals and the departures but does not define the matching uniquely.

Denoting the arrivals by  $\mathbf{a} = (a_1, \dots, a_n)$ , the observed departures by  $\mathbf{d} = (d_1, \dots, d_m)$ , and the set of compatible thinnings by  $S(\mathbf{a}, \mathbf{d})$ , we can write the likelihood of  $\mu$  for fixed  $\mathbf{a}$  in the form

$$L(\mu) = p(\mathbf{d} | \mu, \mathbf{a}) = \sum_{\mathbf{s} \in S(\mathbf{a}, \mathbf{d})} p(\mathbf{d}, \mathbf{s} | \mu, \mathbf{a}) = \sum_{\mathbf{s} \in S(\mathbf{a}, \mathbf{d})} p(\mathbf{d} | \mu, \mathbf{a}, \mathbf{s}) p(\mathbf{s} | \mu, \mathbf{a}).$$

The challenge for the maximum likelihood estimation is the number of compatible thinnings that grows exponentially as a function of the sample size. For instance, in the example of Figure 3(a) with four departures and six arrivals, there are nine compatible thinnings and with additional new events  $a_7$ ,  $a_8$  and  $d_5$ , the number would increase to 23. It seems difficult to draw a random thinning from  $\mathcal{S}(\mathbf{a}, \mathbf{d})$  which would be needed in Monte Carlo approximations of the likelihood. Rejection sampling where random binary sequences are checked for the compatibility fails in practice because the number of possible sequences (choose  $m$  out of  $n$ ) grows even faster than the number of compatible thinnings leading to very low acceptance rates in the sampling. One could consider various sequential sampling techniques where the sampling proceeds iteratively until a compatible thinning is obtained but these approaches seem to lead to incorrect probabilities of compatible thinnings.

Since the maximum likelihood estimator appears to be computationally intractable, we shift the focus to estimators that are based on a certain compatible thinning. When the thinning is fixed, we may use formula (4) to obtain the maximum likelihood estimate given the thinning. This leads to estimators that are in general biased for  $\mu$  but may be useful as initial estimators.

A particular estimator of interest,  $\hat{\mu}_0$ , is a minimum contrast estimator defined as the solution to the minimization problem

$$(\hat{\mu}_0, \hat{s}_1, \dots, \hat{s}_n) = \arg \min_{\mu, s_1, \dots, s_n} -L(\mu, s_1, \dots, s_n | a_1, \dots, a_n, d_1, \dots, d_m), \quad (6)$$

where thinning  $s_1, \dots, s_n$  is required to be compatible with  $a_1, \dots, a_n$  and  $d_1, \dots, d_m$ . As the density function of the exponential distribution decreases as a function of  $x$  for any  $\mu$ , the likelihood is maximized when the lifetimes are as short as possible. This is achieved when the arrival times selected in the thinning are as large as possible. Given the maximizing thinning  $\hat{s}_1, \dots, \hat{s}_n$ , the estimate  $\hat{\mu}_0$  is obtained applying estimator (4) in a modified form

$$\hat{\mu}_0 = \frac{1}{m} \left( \sum_{j=1}^m d_j - \sum_{i=1}^n \hat{s}_i a_i \right). \quad (7)$$

The compatibility requirement for  $\hat{s}_1, \dots, \hat{s}_n$  makes estimator (7) fundamentally different from estimator (5) in two aspects: estimator (7) is always non-negative, and its variance does not increase as a function of the activity time because each departure is implicitly matched with a nearby arrival from the same block. The obtained initial estimator  $\hat{\mu}_0$  underestimates  $\mu$  because the departures leading to the shortest possible lifetimes are chosen in the thinning.

In Section 5, we present an estimation procedure that starts with estimate  $\hat{\mu}_0$  and corrects the bias by means of simulation. We generate datasets with different candidate values of parameter  $\mu$  and calculate the initial estimate  $\hat{\mu}_0(\mu)$  using formula (7). The goal is to find a value of  $\mu$  that produces an initial estimate close to initial estimate  $\hat{\mu}_0$  obtained from the observed data. This specifies an implicit statistical model (Diggle & Gratton, 1984; Hoel & Mitchell, 1971) for the data. To formally justify this approach, we should prove that  $E[\hat{\mu}_0(\mu^{(1)})] < E[\hat{\mu}_0(\mu^{(2)})]$  whenever  $\mu^{(1)} < \mu^{(2)}$  and the arrival rate and the number of arrivals are fixed. It seems hard to show this analytically because the impact of thinning is challenging to evaluate. Therefore, we rely on simulations when studying the properties of the initial estimator and its bias-corrected version.

## 5 | IMPLEMENTATION OF THE PROPOSED ESTIMATION METHOD

Our estimation procedure consists of five steps: (1) dividing the timeline into blocks (Section 5.1), (2) thinning the arrivals block-wise (Section 5.2), (3) applying estimator (7) (Section 5.2), (4) using noisy binary search for bias (Section 5.3) and (5) estimating the standard error by using bootstrap techniques (Section 5.4). The R code for the estimation procedure is available for download at <https://github.com/JuhaKarvanen/radioactive>.

### 5.1 | Formation of blocks

As the first step of the proposed estimation procedure, we divide the timeline into blocks on the basis of observed arrival and departure times. We work with a sequence  $v_1, \dots, v_{n+m}$  of events where either  $v_i = a$  (arrival) or  $v_i = d$  (departure). The number of arrivals is greater than or equal to the number of departures. The sequence must start with an arrival and end with a departure. The blocks are constructed processing the events sequentially from the end  $v_{n+m}$  to the beginning  $v_1$ . The numbers of arrivals and departures in the current block are tracked cumulatively during the processing. When the following three conditions are met in this process, a cut-point between the current block and the next block is set between events  $v_{i-1}$  and  $v_i$ :

- (a) event  $v_i$  is an arrival,
- (b) event  $v_{i-1}$  is a departure, and
- (c) the current number of arrivals in the block that is currently constructed is greater than the current number of departures.

The next block starts with departure  $v_{i-1}$  and the counters for the cumulative numbers of arrivals and departures are set to zero.

Let us consider an example sequence of ten events

$$(v_1, \dots, v_{10}) = (a_1, d_1, a_2, a_3, d_2, a_4, a_5, a_6, d_3, d_4)$$

shown in Figure 3. When events  $v_5 = d_2$  and  $v_6 = a_4$  are processed, we have a situation where event  $v_6$  is an arrival, event  $v_5$  is a departure, the number of arrivals is 3 and the number of departures is 2. Thus all three conditions are met and the cut-point is set between events  $v_5$  and  $v_6$ . Proceeding in the same manner we end up having three blocks, which are  $(a_1, d_1)$ ,  $(a_2, a_3, d_2)$  and  $(a_4, a_5, a_6, d_3, d_4)$ .

### 5.2 | Thinning of arrivals in blocks

In the second step, some of the arrivals are removed to restore the balance between arrivals and departures. We have chosen to use the thinning that provides the solution to the minimization problem (6). If a block has  $A$  arrivals and  $D$  departures, the  $A-D$  smallest arrival times are removed. In the example of Figure 3, the blocks after thinning would be  $(a_1, d_1)$ ,  $(a_3, d_2)$  and  $(a_5, a_6, d_3, d_4)$ . From the formed blocks, the initial estimate (7) is calculated.

The division into blocks is needed to ensure the compatibility of the thinning. Sometimes ‘wrong’ arrivals are removed in the thinning, which causes bias in the initial estimate  $\hat{\mu}_0$ . For instance, in the example shown in Figure 3, arrivals  $a_2$  and  $a_4$  are removed, even if they could be the ones that match with the observed departures.

A referee proposed an alternative approach that will lead to the same initial estimate as the formation of blocks and the thinning described above. In this approach, departures are processed starting from the end and each departure time  $d_j$  is matched with the largest possible arrival time  $a_i$ ,  $a_i < d_j$ , that has not been matched yet. After matching estimator (2) can be used to obtain the same initial estimate  $\hat{\mu}_0$  as described above. This matching is illustrated in Figure 2 where we see that the distribution of the obtained pseudo-lifetimes clearly differs from the exponential distribution fitted to the original data. Conceptually, we prefer the block-based definition because it does not require explicit matching and keeps the formation of the blocks and the thinning separate from each other.

### 5.3 | Bias correction

In order to find an unbiased estimator for the mean lifetime  $\mu$  we apply an implicit statistical model and simulation based estimation. The initial estimate  $\hat{\mu}_0$  obtained from the thinned data by estimator (7) provides a starting point. We simulate data with different values of  $\mu$ , apply thinning as described above, and estimate the mean lifetime with estimator (7). We aim to find a value of  $\mu$  that produces an estimate sufficiently close to  $\hat{\mu}_0$ . Binary search is applicable for this task because the parameter of interest is one-dimensional. However, we must apply noisy binary search (Ben-Or & Hassidim, 2008; Karp & Kleinberg, 2007; Rivest et al., 1980) that takes sampling variation into account. In practice, this means that several datasets must be generated to ensure that the mean of the obtained estimates is close to  $\hat{\mu}_0$  with a high confidence. The detailed procedure for the bias correction is given in Algorithm 1 whose steps and notations are described below.

In Algorithm 1, we use the following notations. The candidate for  $\mu$ , denoted by  $M$ , can have values in the interval  $(L, U)$  initialized as  $(0, \infty)$ . Variable  $M_0$  stands for the average of estimates (7) calculated (line 8) from samples simulated with an exponential distribution with mean  $M$  (lines 5–7). For each  $M$ , we simulate at least  $K_{\min}$  samples (line 4), and at maximum  $K_{\max}$  samples (line 13). The distance between  $M_0$  and the estimated value  $\hat{\mu}_0$  is denoted by  $D$ . The goal is to find the value of  $M$  that gives the smallest  $D$ . The value  $M$  is accepted and returned (line 10) if  $|D| \leq \delta \hat{\mu}_0$ , where  $\delta$  is the tolerance given by the user. The maximum number of steps, that is, the maximum number of  $M$  candidates considered, is  $J_{\max}$  (line 22).

If the candidate  $M$  is not accepted on the basis of  $K_{\min}$  samples, the confidence interval  $(D_L, D_U)$  for  $E[D]$  is calculated (line 11)  $D \pm z_{1-\alpha/2} \text{SE}(M_0)$ , where  $\text{SE}(M_0)$  is the standard error of  $M_0$  estimated from the  $K_{\min}$  samples and  $z_{1-\alpha/2}$  is the  $1 - \alpha/2$  quantile of the standard normal distribution. If the confidence interval includes zero, the loop (line 13) continues simulating the samples with the current  $M$  until  $M$  is accepted (line 19) or  $K_{\max}$  is achieved (line 13) or the recalculated confidence interval does not include zero (line 13).

In the latter two cases, if the number of  $M$  candidates  $J_{\max}$  is not reached (line 22), the lower or upper limit of  $M$  is first updated (lines 24–25 or lines 27–28). The new value for  $M$  is then calculated in two ways depending on the value of  $U$ . If  $U < \infty$  (lines 26 and 29) or if  $U$  is  $\infty$  and

$D > 0$  (lines 23 and 25), the value of  $M$  is the average of  $L$  and  $U$ . If  $U$  is  $\infty$  and  $D < 0$  (line 24), the current  $M$  is doubled.

Choosing large  $K_{\min}$ ,  $K_{\max}$  and  $J_{\max}$  and small tolerance  $\delta$  leads to more precise estimates but requires more computational time. This trade-off has practical importance in simulation experiments where the analysis is repeated thousands of times.

## 5.4 | Parametric bootstrap

We apply parametric bootstrap (Efron & Tibshirani, 1993) to obtain the standard error and confidence intervals for the bias-corrected estimate of the parameter  $\mu$ . When simulating bootstrap samples, the lifetimes are generated from the exponential distribution using the bias-corrected estimate as the mean lifetime. The arrival times are fixed to the arrival times in the data. The departure times are obtained by adding the lifetimes to the arrival times and each departure is removed with probability 0.5. For each bootstrap sample, the proposed estimation method with four steps (forming blocks, thinning, estimating initial value and using a bias correction) is applied in order to obtain the bias-corrected estimate. The standard error is estimated by the standard deviation of the bias-corrected estimates and the confidence intervals are calculated with the percentile method.

## 6 | SIMULATION EXPERIMENTS

Next we study the properties of the proposed estimator in simulation experiments. In general, simulations may not lead to sufficient understanding on the performance of a new method because it is difficult to ensure that all relevant setups are considered. Here, simulations provide relatively strong evidence because all setups can be essentially characterized with only two quantities: the number of events which depends linearly on the activity time, and the ratio of the arrival rate  $\lambda_a$  and the decay rate  $\lambda$ . A wide range of values for these quantities can be easily covered in simulations. In addition, Algorithm 1 has control parameters for the numeric accuracy.

The arrivals follow a Poisson process with known arrival rate  $\lambda_a$  and the probability of detecting each departure is 0.5. The mean lifetime  $\mu = 1/\lambda$  is fixed to 1 without loss of generality and the arrival rate is altered. In all simulations, the control parameters of Algorithm 1 have values: tolerance  $\delta = 0.001$ , risk level  $\alpha = 0.05$ , minimum sample size  $K_{\min} = 5$ , maximum sample size  $K_{\max} = 100$  and maximum number of iterations  $J_{\max} = 100$ .

### 6.1 | Simulations with fixed number of arrivals

In the first simulation, reported in Table 1 and in Figure 4, the activity time is chosen so that the expected number of arrivals during the activity period is always 2000. It is seen from Table 1 that the initial estimates are often strongly biased but the bias correction technique described in Section 5.3 reduces the bias to a few per cent or less in all setups. The biases and standard errors are calculated from 10,000 simulation runs per each arrival rate.

We consider the limiting behaviour of the estimators understanding that the setups with very small or very large arrival rate may not be implementable in real experiments. In the setups of

**Algorithm 1** Algorithm for bias correction.

**Input:** Initial estimate  $\hat{\mu}_0$ , arrival times  $a_1, \dots, a_n$ , control parameters: tolerance  $\delta$ , risk level  $\alpha$ , minimum sample size  $K_{\min}$ , maximum sample size  $K_{\max}$ , maximum number of iterations  $J_{\max}$

**Output:** A bias-corrected estimate  $\hat{\mu}$  for the mean lifetime

- 1: **Let**  $M = 2\hat{\mu}_0$ ,  $L = 0$ ,  $U = \infty$ ,  $j = 0$
- 2: **Repeat**
- 3:   **Let**  $j = j + 1$
- 4:   **Repeat** for  $K$  in 1 to  $K_{\min}$ :
- 5:     Simulate lifetimes  $t_1, \dots, t_n$  from  $\text{Exp}(M^{-1})$ , calculate departure times  $a_1 + t_1, \dots, a_n + t_n$ , and generate  $n$  random variables from  $\text{Bern}(0.5)$  to find out whether each departure is observed or not.
- 6:     Form the blocks as described in Section 5.1 and apply the thinning of Section 5.2 for the arrivals and the observed departures.
- 7:     Apply estimator (7) for the thinned data to obtain estimate  $\hat{\mu}_K$ .
- 8:     **Let**  $M_0 = \frac{1}{K_{\min}} \sum_{k=1}^{K_{\min}} \hat{\mu}_k$
- 9:     **Let**  $D = M_0 - \hat{\mu}_0$
- 10:     **If**  $|D| \leq \delta \hat{\mu}_0$  **then return**  $M$
- 11:     Compute the  $1 - \alpha$  confidence interval  $(D_L, D_U)$  for  $E[D]$ .
- 12:     **Let**  $K = K_{\min} + 1$
- 13:     **While**  $(K < K_{\max}$  and  $D_L < 0$  and  $D_U > 0)$  **do**
- 14:       Simulate lifetimes  $t_1, \dots, t_n$  from  $\text{Exp}(M^{-1})$ , calculate departure times  $a_1 + t_1, \dots, a_n + t_n$ , and generate  $n$  random variables from  $\text{Bern}(0.5)$  to find out whether each departure is observed or not.
- 15:       Form the blocks as described in Section 5.1 and apply the thinning of Section 5.2 for the arrivals and the observed departures.
- 16:       Apply estimator (7) for the thinned data to obtain estimate  $\hat{\mu}_K$ .
- 17:       **Let**  $M_0 = \frac{1}{K} \sum_{k=1}^K \hat{\mu}_k$
- 18:       **Let**  $D = M_0 - \hat{\mu}_0$
- 19:       **If**  $|D| \leq \delta \hat{\mu}_0$  **then return**  $M$
- 20:       Compute the  $1 - \alpha$  confidence interval  $[D_L, D_U]$  for  $E[D]$ .
- 21:       **Let**  $K = K + 1$
- 22:     **If**  $j \geq J_{\max}$  **then return**  $M$
- 23:     **If**  $U = \infty$  **then**
- 24:       **If**  $D < 0$  **then let**  $L = M$ ,  $M = 2M$
- 25:       **If**  $D > 0$  **then let**  $U = M$ ,  $M = (U + L)/2$
- 26:     **else**
- 27:       **If**  $D < 0$  **then let**  $L = M$
- 28:       **If**  $D > 0$  **then let**  $U = M$
- 29:       **Let**  $M = (U + L)/2$

the first simulation, the initial estimate is unbiased when the arrival rate approaches to zero or infinity. When the arrival rate approaches zero, a departure (observed or unobserved) occurs almost always before the next arrival occurs. This means that with a high probability a block will contain only one departure and this departure and the arrival preceding it correspond to each other. In the thinning, other arrivals in the block (if any) are removed and only the departure and its corresponding arrival will remain. When the arrival rate approaches infinity and the expected number of arrivals is fixed, all arrivals occur in a short interval and almost all departures occur after this interval. The thinning causes only a small bias because all arrival times are almost the same. Consequently, the situation is close to an ideal experiment where all arrivals occur at the same time and the departures are observed without interference. The limiting standard error  $1/\sqrt{1000} \approx 0.032$  is obtained by Equation (3) where  $n = 1000$  equals the number of observed departures (half of 2000 arrivals).

When we look at the results as a function of arrival rate  $\lambda_a$  keeping in mind that the number of arrivals was fixed to be 2000, there are two effects in the opposite directions. When the arrival rate increases, it becomes more difficult to match the arrivals and the departures and the bias caused by thinning increases. When the activity time becomes shorter, more departures occur

**TABLE 1** The bias and the standard error of the estimates of the mean lifetime as functions of the arrival rate when the expected number of arrivals is fixed to 2000

$\lambda_a$	Activity time	Average number of departures during + after	Initial estimate $\hat{\mu}_0$		Corrected estimate $\hat{\mu}$	
			Bias ( $10^{-6}$ )	SE ( $10^{-6}$ )	Bias ( $10^{-6}$ )	SE ( $10^{-6}$ )
0.0010	2,000,000	1002 + 0.00100	-1492	32,662	-1004	33,382
0.010	200,000	1001 + 0.00600	-4999	31,380	-965	32,411
0.10	20,000	1001 + 0.0470	-48,541	29,372	770	33,844
0.20	10,000	1000 + 0.111	-91,949	29,051	1618	36,708
0.50	4000	999 + 0.265	-206,872	24,906	1949	41,577
1.0	2000	998 + 0.523	-350,429	22,350	1759	55,555
1.5	1333	999 + 0.767	-451,435	20,402	3082	70,196
2.0	1000	1000 + 0.986	-525,394	17,839	4550	80,955
5.0	400	996 + 2.42	-731,918	12,721	17,705	152,324
10	200	994 + 5.04	-838,022	9392	16,988	196,961
15	133	993 + 7.53	-878,758	8617	16,886	209,075
20	100	989 + 10.0	-899,805	8256	-83	199,785
50	40	978 + 24.9	-926,573	10,292	-4583	130,901
100	20	951 + 49.9	-907,225	13,729	-4372	91,541
1000	2	568 + 433	-406,097	31,927	1708	35,541
10,000	0.2	93.8 + 906	-49,995	29,631	518	30,759
100,000	0.02	9.95 + 988	-5715	31,109	-1683	32,205

*Notes:* For each setup, the reported statistics are based on 10,000 simulation runs. The average bias (in units of  $10^{-6}$ ) was calculated as the average of the differences of the 10,000 estimates and the true mean lifetime  $\mu = 1$ . The standard errors (SE, in units of  $10^{-6}$ ) were calculated as the standard deviations of the 10,000 estimates. The third column reports the average number of observed departures during the activity time and after the activity time.

after the activity period (Table 1, third column). The expected number of these departures can be calculated with the formula (see Appendix for the derivation)

$$\lambda_a(\mu - e^{-t/\mu})/2, \quad (8)$$

where  $t$  is the activity time. The departures at the post-activity period are highly informative because it is known for sure that the corresponding arrivals occurred during the activity period. In Table 1, the initial estimate has the largest bias when  $\lambda_a = 50$ . In this case, the average number of the post-activity departures is still quite small (24.9 departures) but the bias caused by thinning is already high.

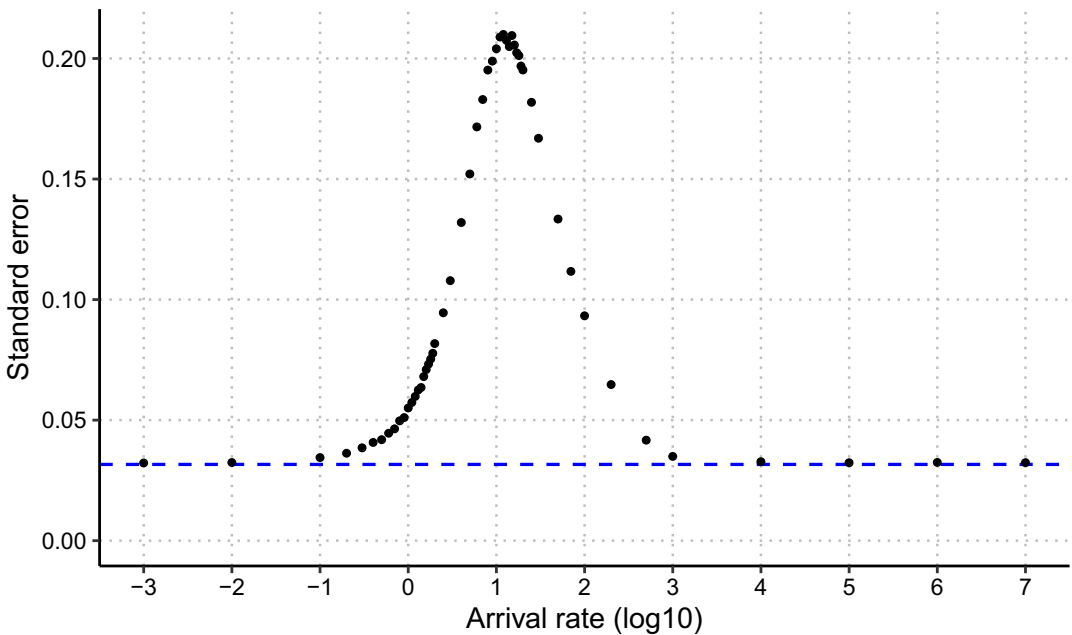
The bias of the initial estimate, the standard error of the initial estimate and the standard error of the corrected estimate are closely related. When the bias is large, initial estimate  $\hat{\mu}_0$  is small and its standard error is also small because it is proportional to the mean (Equation 3). The large bias means that it is more difficult to correct the bias by Algorithm 1 and these difficulties are reflected in the standard error of the corrected estimate. In Table 1, the highest standard error of the corrected estimate occurs at  $\lambda_a = 15$  when the initial estimate has a large bias and small

standard error. The uncertainty of the corrected estimate is the price we pay for the bias in the initial estimate.

The results give indirect information on the relative efficiency of the proposed estimation method compared with the maximum likelihood estimation. The optimal method (maximum likelihood) cannot perform better than the theoretical lower limit (standard error 0.032). We see that if the arrival rate  $\lambda_a$  is smaller than 0.5 or larger than 500, the optimal method cannot be much better than the proposed method. When  $\lambda_a$  is between 0.5 and 500, there may be room for improvement. The maximum standard error 0.209 was obtained when  $\lambda_a = 15$ . This means that in the imaginary extreme case, the optimal method might be  $0.210/0.032 = 6.5$  times more efficient. In other words, we have shown that the practical gain from a more efficient estimation method could be considerable but not huge.

## 6.2 | Simulations with fixed activity time

In the second simulation, reported in Table 2 and in Figure 5, the activity time is fixed to a value (Table 2: 2000; Figure 5: 1000, 2000 or 3000) but the arrival rate is varied from 0.1 to 1000. In this setup, the smallest standard error of the bias-corrected estimate is obtained when the arrival rate is very large. Practically all information comes then from departures observed after the activation period. Equation (8) tells us that for long arrival times, the expected number of these departures is approximately  $\lambda_a \mu / 2$  which does not depend on the activity time. For this reason, the curves for different activity times coincide for large arrival rates in Figure 5.



**FIGURE 4** The standard error for the bias-corrected estimate of the mean lifetime as function of the logarithm of the arrival rate in the first simulation experiment. The expected number of arrivals is always 2000 and the true mean lifetime is 1. The dashed line shows the theoretical lower limit for the standard error. The results are based on 10,000 simulation runs per each arrival rate



**TABLE 2** The bias and the standard error of the estimates of the mean lifetime as functions of arrival rate  $\lambda_a$  in the first simulation experiment where the activity time is fixed to 2000

$\lambda_a$	Activity time	Average number of departures during + after	Initial estimate $\hat{\mu}_0$		Corrected estimate $\hat{\mu}$	
			Bias ( $10^{-6}$ )	SE ( $10^{-6}$ )	Bias ( $10^{-6}$ )	SE ( $10^{-6}$ )
0.10	2000	99.9 + 0.0489	-48,011	95,557	1396	108,248
0.20	2000	200 + 0.104	-92,499	65,112	2933	80,925
0.50	2000	500 + 0.246	-208,278	36,598	-135	59,306
1.0	2000	1000 + 0.505	-351,339	22,336	340	54,679
2.0	2000	1999 + 0.987	-525,862	12,994	1454	57,994
5.0	2000	4997 + 2.50	-734,770	5577	4132	69,812
10	2000	9993 + 4.99	-844,088	2707	5778	84,146
20	2000	19,990 + 10.0	-913,106	1244	9171	100,305
30	2000	29,985 + 15.0	-939,367	763	7560	105,924
40	2000	39,978 + 20.0	-953,285	553	5878	108,452
50	2000	49,977 + 25.0	-961,929	429	6329	108,522
60	2000	59,972 + 30.1	-967,830	348	4904	104,692
70	2000	69,967 + 34.9	-972,118	300	2764	103,424
80	2000	79,962 + 39.9	-975,369	255	2699	97,856
90	2000	89,954 + 45.0	-977,923	228	1894	95,142
100	2000	99,952 + 50.0	-979,984	206	537	91,644
200	2000	199,900 + 100	-989,444	116	-11	67,503
500	2000	499,748 + 250	-995,285	67	-800	43,120
1000	2000	999,498 + 500	-997,259	47	86	30,753

*Notes:* For each setup, the reported statistics are based on 5000 simulation runs. The average bias (in units of  $10^{-6}$ ) was calculated as the average of the differences of the 5000 estimates and the true mean lifetime  $\mu = 1$ . The standard errors (SEs, in units of  $10^{-6}$ ) were calculated as the standard deviations of the 5000 estimates. The third column reports the mean number of observed departures during the activity time and after the activity time.

When we look at the results as a function of arrival rate  $\lambda_a$ , there are again effects in opposite directions. When the arrival rate increases, the number of arrivals increases, which as such decreases the uncertainty of the initial estimate (Table 2, fifth column). At the same time, it becomes more difficult to match the arrivals and the departures, which increases the bias of the initial estimate (Table 2, fourth column). The high arrival rate also means that the number departures after the activity period is high. Using Equation (8) we learn that the ratio of the expected number the post-activity departures and the expected number of all departures

$$\frac{\lambda_a(\mu - e^{-t/\mu})}{\lambda_a t} = \frac{\mu - e^{-t/\mu}}{t} \approx \frac{\mu}{t}$$

does not depend on the arrival rate. In Table 2, the expected proportion of post-activity departures is thus  $1/2000$ .

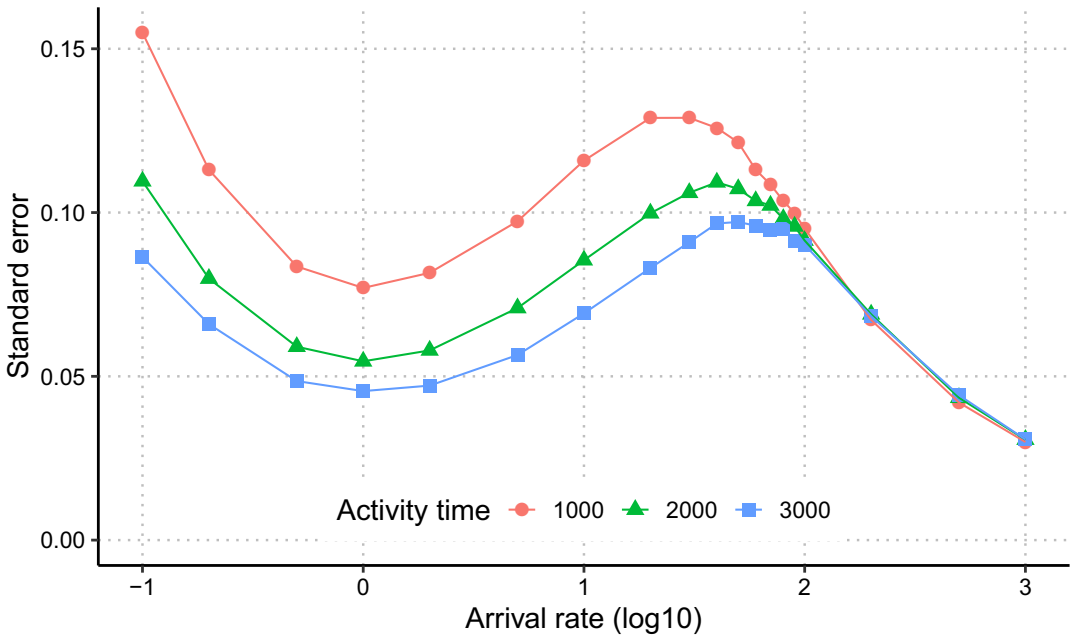


FIGURE 5 The standard error for bias-corrected estimate of the mean lifetime as function of the logarithm of the arrival rate. The activity time is 1000, 2000 or 3000 and the true mean lifetime is 1. The results are based on 5000 simulation runs per each combination of arrival rate and activity time

As the composition of the different effects, we see the non-trivial behaviour illustrated in Figure 5. There seems to be a local minimum at  $\lambda_a = 1$ , that is, when the arrival rate and the decay rate are equal. The location of a local maximum varies as a function of activity time. For activity times 1000, 2000 and 3000 the local maxima are located around  $\lambda_a = 30$ ,  $\lambda_a = 40$  and  $\lambda_a = 50$  respectively.

### 6.3 | Simulations studying the effect of sample size

In the third simulation, reported in Table 3 and in Figure 6, we study how the standard error of the estimated mean lifetime behaves as a function of the expected number of arrivals  $\lambda_a t$ . It is seen from Figure 6 that for low arrival rates,  $\lambda_a = 0.1$  and  $\lambda_a = 1$ , the relation of the standard error and the expected number of arrivals is ordinary, that is, approximately linear in the log–log scale. When  $\lambda_a = 1$ , the standard errors are about 1.7 times the standard errors (the dashed line) one would obtain if the arrivals and the observed departures were matched.

For high arrival rates,  $\lambda_a = 100$  and  $\lambda_a = 500$ , increasing the expected number of arrivals does not decrease the standard error after a saturation point, as also seen in Figure 5. In these cases, the information on  $\mu$  comes mainly from the departures observed after the activity period. Increasing the activity time has only a negligible effect on the number of these departures if the activity time is already long because it is unlikely that an arrival observed at the beginning of the activity period matches with a departure observed after the activity period. For intermediate arrival rates,  $\lambda_a = 5$  and  $\lambda_a = 20$ , the curves in Figure 6 are concave but seem to approach linear behaviour when the expected number of arrivals increases.

**TABLE 3** The bias and the standard error of the estimates of the mean lifetime as functions of expected number of arrivals for various arrival rates  $\lambda_a$

$\lambda_a$	Activity time	Expected number of arrivals	Average number of departures during + after	Initial estimate $\hat{\mu}_0$		Corrected estimate $\hat{\mu}$	
				Bias ( $10^{-6}$ )	SE ( $10^{-6}$ )	Bias ( $10^{-6}$ )	SE ( $10^{-6}$ )
0.10	4000	400	200 + 0.0483	-47,825	67,157	1688	76,207
0.10	25,000	2500	1251 + 0.0574	-48,482	27,007	524	31,065
0.10	100,000	10,000	5001 + 0.0469	-48,381	13,313	-292	15,450
1.0	400	400	200 + 0.489	-349,507	50,985	5926	123,255
1.0	2500	2500	1249 + 0.497	-351,147	20,035	200	49,474
1.0	10,000	10,000	4999 + 0.517	-351,081	9779	558	24,331
5.0	80	400	197 + 2.50	-720,498	31,055	26,758	272,885
5.0	500	2500	1247 + 2.49	-733,237	11,083	8839	133,990
5.0	2000	10,000	4998 + 2.45	-734,854	5559	3152	70,350
20	20	400	190 + 10.0	-843,705	32,663	-10,137	203,248
20	125	2500	1240 + 9.97	-902,892	6704	-4563	188,801
20	500	10,000	4988 + 10.1	-911,069	2685	6415	155,729
100	4	400	151 + 49.1	-624,596	62,401	-1373	92,494
100	25	2500	1200 + 49.9	-921,792	11,557	-3746	95,091
100	100	10,000	4950 + 49.9	-966,021	2993	-4565	96,269
500	0.8	400	62.5 + 138	-184,038	70,613	1994	72,859
500	5	2500	1001 + 248	-700,996	23,783	-570	41,521
500	20	10,000	4748 + 250	-921,150	6311	315	41,305

*Notes:* For each setup, the reported statistics are based on 5000 simulation runs. The average bias (in units of  $10^{-6}$ ) was calculated as the average of the differences of the 5000 estimates and the true mean lifetime  $\mu = 1$ . The standard errors (SE, in units of  $10^{-6}$ ) were calculated as the standard deviations of the 5000 estimates. The third column reports the mean number of observed departures during the activity time and after the activity time.

## 7 | EXAMPLES BASED ON $^{155}\text{Lu}$ ALPHA DECAY DATA

We apply the methods of Section 3 to Datasets 1 and 2, and the method of Sections 4 and 5 to Datasets 3A–F and 4A–4F. Table 4 summarizes the estimates and their confidence intervals. Original data (Dataset 1) and its unmatched densified variant (Dataset 2) give the same estimates for reasons explained in Section 3.2. An estimate for the half-life 2.61 ms (2.57,2.65) is obtained by multiplying the mean lifetime by  $\log(2)$ . This estimate is close to the estimates reported in the literature (Page et al., 1996).

For incomplete data (Datasets 3A–3F), the point estimates are not exactly the same but still relatively close to the estimate obtained from Dataset 1. The uncertainty of the estimates behaves similarly to the first simulation experiment. (The standard errors and the arrival rates should be divided by the mean lifetime and accounted for the different sample sizes before they can be compared with those in Table 1.) When the activity time decreases and the arrival rate increases, the bias of the initial estimate increases. In Table 4 this is seen as an increase in the standard error

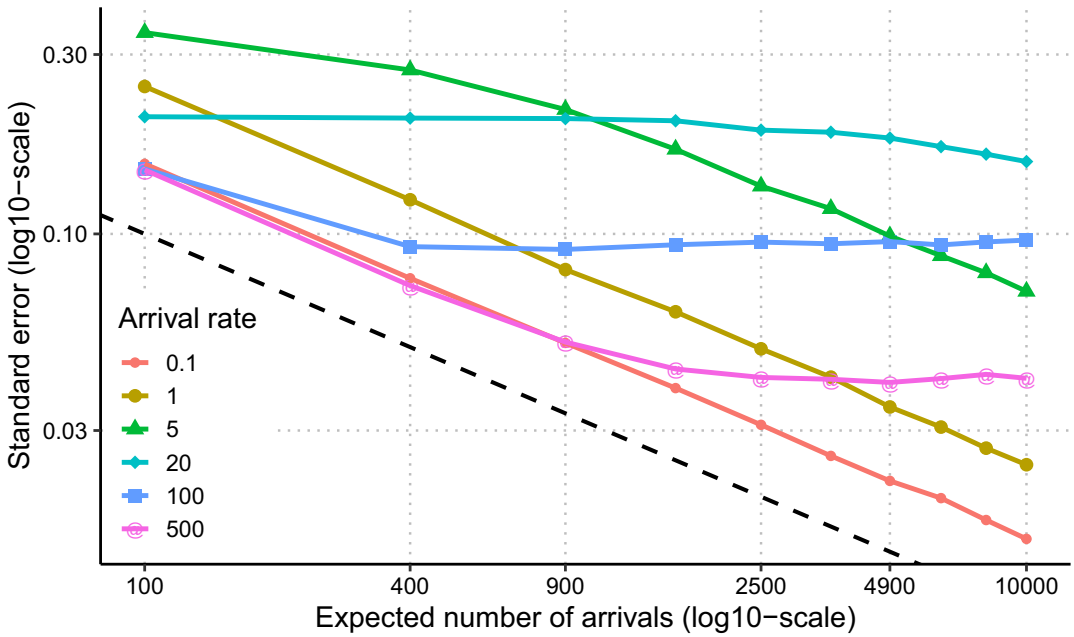


FIGURE 6 The standard error for bias-corrected estimate of the mean lifetime as function of the expected number of arrivals. The reference line (dashed) has the slope  $-0.5$  in the log scale, which corresponds to the factor  $1/\sqrt{n}$  in the formula of the standard error. The results are based on 5000 simulation runs per each combination of arrival rate and expected number of arrivals

TABLE 4 The estimates of the mean lifetime from the datasets 1, 2, 3A–3F and 4A–4F

Dataset	Activity time	Arrival rate (events/ms)	No. of departures during + after	Estimate (SE) (ms)	95% confidence interval (ms)
1	14.6 h	$3.06 \times 10^{-4}$	16,068 + 1	3.760 (0.030)	(3.702, 3.818)
2	53.0 s	$3.06 \times 10^{-1}$	16,067 + 2	3.760 (0.030)	(3.702, 3.818)
3A	1.46 h	$3.06 \times 10^{-3}$	8048 + 0	3.760 (0.046)	(3.671, 3.850)
3B	8.83 min	$3.06 \times 10^{-2}$	8048 + 0	3.733 (0.046)	(3.644, 3.826)
3C	53.0 s	$3.06 \times 10^{-1}$	8047 + 1	3.808 (0.079)	(3.647, 3.953)
3D	5.30 s	$3.06 \times 10^0$	8045 + 3	3.334 (0.327)	(2.731, 4.038)
3E	530 ms	$3.06 \times 10^1$	7992 + 56	3.726 (0.380)	(3.039, 4.409)
3F	53 ms	$3.06 \times 10^2$	7485 + 563	3.602 (0.103)	(3.399, 3.805)
4A	1.46 h	$3.06 \times 10^{-3}$	3992 + 0	3.809 (0.063)	(3.686, 3.935)
4B	8.83 min	$3.06 \times 10^{-2}$	3992 + 0	3.799 (0.065)	(3.684, 3.939)
4C	53.0 s	$3.06 \times 10^{-1}$	3992 + 0	4.003 (0.135)	(3.734, 4.237)
4D	5.30 s	$3.06 \times 10^0$	3991 + 1	3.625 (0.720)	(2.739, 5.583)
4E	530 ms	$3.06 \times 10^1$	3971 + 21	4.040 (0.807)	(2.671, 5.902)
4F	53 ms	$3.06 \times 10^2$	3703 + 289	3.855 (0.159)	(3.550, 4.177)

Notes: The number of arrivals is 16,069 for all datasets. The confidence intervals for Datasets 1 and 2 are obtained by using the asymptotic properties of the maximum likelihood approach. For Datasets 3A–3F and 4A–4F, the standard errors and confidence intervals were calculated using parametric bootstrap with 1000 replications.

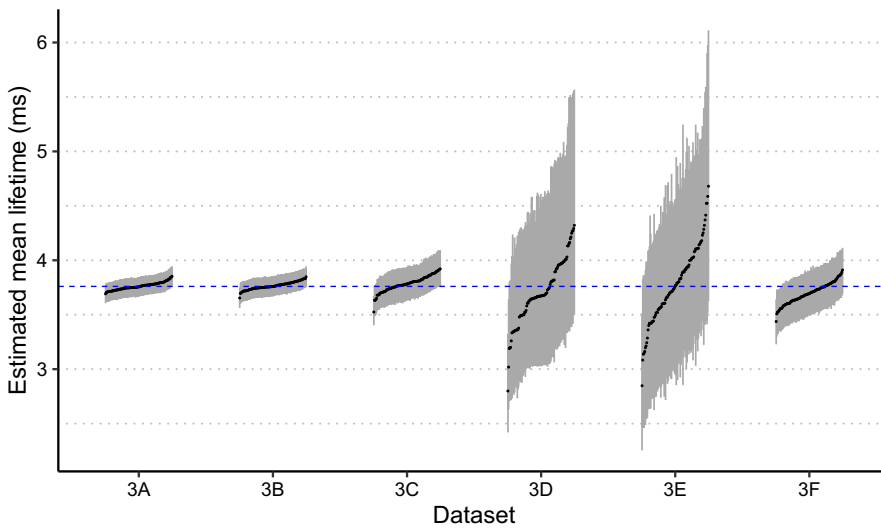


FIGURE 7 The point estimates (dots) and their 95% confidence intervals (vertical lines) for Datasets 3A–3F with 100 independent repetitions of random removal of departures. For each dataset, the results are ordered by the point estimate in increasing order from left to right. The horizontal dashed line shows the point estimate from Dataset 1

of the bias-corrected estimate, especially in Datasets 3D and 3E. However, increasing arrival rate leads to higher number of departures after the activity period. Due to the departures after the activity period, the standard error is smaller for Dataset 3F than for Datasets 3D and 3E. It is remarkable that the activity time of 53 ms (Dataset 3F) gave the standard error 0.10 which is only twice the standard error with the activity time of 1.46 h (Dataset 3A).

The results for Datasets 4A–4F show that the method works even if the probability for observing a departure is only 0.25. We see that the ratio of the standard errors for Datasets 3A and 4A, 1.37, is close to  $\sqrt{2}$  as one would expect when the sample size is halved. However, the same ratio for Datasets 3D and 4D is 2.2 indicating that in this setup the estimation becomes relatively more difficult when the probability for observing a departure drops from 0.5 to 0.25.

The results demonstrate that the additional loss of information due to the thinning is not necessarily large. The standard errors for Datasets 1 (16,069 departures) and 3A (8042 departures) are 0.030 and 0.046 respectively. If we scale the standard error of Dataset 1 from sample size 16,069 to sample size 8042, we obtain 0.042 which is 9% smaller than 0.046.

As an additional analysis, we repeated the creation of Datasets 3A–3F 100 times (regenerating the random removal of departures) and used the same estimation procedure. Figure 7 shows the obtained point estimates and their 95% confidence intervals. The results support the conclusion that even if the uncertainty of the estimate depends on the arrival rate, the proposed estimation method works reliably for a wide range of arrival rates.

## 8 | DISCUSSION

We have proposed a method for the estimation of the mean lifetime in situations where the arrivals and the departures are unmatched and the departures are missing completely at random. The method is derived for experiments in nuclear physics where the goal is to

characterize particles by their expected lifetime. The estimation relies on the exponential distribution of lifetimes—an assumption that follows from the laws of physics. Due to the memoryless property of the exponential distribution, the information on the matching of arrivals and departures is not needed in the estimation. In principle, the method could be applicable to other problems as well if the assumption of the exponential distribution can be justified.

Simulations and real data examples demonstrate that the mean lifetime can be successfully estimated from partially observed departures. The estimation method combines an initial estimator and a numerical bias correction technique. As the exponential distribution has only one parameter, the noisy binary search can be applied to find an unbiased estimator in the minimum contrast estimation. One could apply the same bias correction technique with some other initial estimator that is based on a different thinning strategy. The probability of observing a departure was known in our examples but could be estimated from the data as well.

We focused on the problems caused by partially observed departures and did not consider misclassification of arrivals in this paper. Further challenges arise in experiments where a large proportion of the arrivals cannot be reliably separated from the background (as defined in Section 2.1). There is also a risk in some experiments that some background events are mislabelled as departures. Whether good estimators can be derived for these situations is an open question.

The simulation results reveal an interesting relation between the arrival rate and standard error of the estimate. If the activity time is fixed (the setup of the second simulation experiment), the arrival rate should be either equal to the decay rate or very large. The latter case, however, has some limitations: First, very large arrival rates are usually practically unreachable in real-world experiments. Second, it is crucial that the follow-up period is longer than the activity period. Third, increasing the activity time may have only very small impact on the accuracy of the estimates after a saturation point has been reached as seen from Figure 6. Therefore, if any prior information on the mean lifetime is available, it seems that an optimal design should have an arrival rate equal to the expected decay rate. The theoretical justification for this empirical observation remains as an open problem.

In nuclear physics, the estimation of the mean lifetime for a particle may be only one of multiple goals in an experiment. This means that the arrival rate might be far from the optimal for a particular estimation task. Fortunately, the proposed estimation method can be applied with a wide range of arrival rates. This will enable using the data from all DSSD strips. Compared to approaches that utilize only the strips where the arrival rate is sufficiently low for the matching of arrivals and departures, the proposed estimation method easily multiplies the amount of available data. As a consequence, the estimates will be more accurate or the required accuracy can be achieved in a shorter experiment.

## ACKNOWLEDGEMENTS

The authors thank Antti Penttinen for useful discussions and the referees for constructive comments. The associate editor is acknowledged for proposing the simulation experiment reported in Section 6.3. CSC – IT Center for Science, Finland, is acknowledged for computational resources.

## ORCID

Juha Karvanen  <http://orcid.org/0000-0001-5530-769X>

## REFERENCES

- Ben-Or, M. & Hassidim, A. (2008) The Bayesian learner is optimal for noisy binary search (and pretty good for quantum as well). In: *2008 49th annual IEEE symposium on foundations of computer science*, IEEE. pp. 221–230.
- Cox, D.R. & Lewis, P.A.W. (1966) *The statistical analysis of series of events*. London: Methuen & Co.
- Diggle, P.J. & Gratton, R.J. (1984) Monte Carlo methods of inference for implicit statistical models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 46(2), 193–212.
- Efron, B. & Tibshirani, R.J. (1993) *An introduction to the bootstrap*. Boca Raton, FL: CRC Press.
- Geiger, H. & Nuttall, J. (1911) Lvii. The ranges of the  $\alpha$  particles from various radioactive substances and a relation between range and period of transformation. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 22(130), 613–621.
- Herzåän, A., Juutinen, S., Auranen, K., Grahn, T., Greenlees, P.T., Hauschild, K. et al. (2015) Detailed spectroscopy of  $^{193}\text{Bi}$ . *Physical Review C*, 92, 044310.
- Hoel, D.G. & Mitchell, T.J. (1971) The simulation, fitting, and testing of a stochastic cellular proliferation model. *Biometrics*, 27(1), 191–199.
- Karp, R.M. & Kleinberg, R. (2007) Noisy binary search and its applications. In: *Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms*. Society for Industrial and Applied Mathematics, pp. 881–890.
- Lazarus, I., Appelbe, E.E., Butler, P.A., Coleman-Smith, P.J., Cresswell, J.R., Freeman, S.J. et al. (2001) The GREAT triggerless total data readout method. *IEEE Transactions on Nuclear Science*, 48(3), 567–569.
- Liukkonen, E. (1993) New K130 cyclotron at Jyväskylä. In: *Cyclotrons and their applications: proceedings of the 13th international conference, Vancouver, 1992*, Singapore. World Scientific, pp. 22–27.
- Lyons, L. (2008) Open statistical issues in particle physics. *The Annals of Applied Statistics*, 2(3), 887–915.
- Page, R.D., Woods, P.J., Cunningham, R.A., Davinson, T., Davis, N.J., James, A.N. et al. (1996) Radioactivity of neutron deficient isotopes in the region  $n > 82 > z$ . *Physical Review C*, 53, 660–670.
- Page, R., Andreyev, A., Appelbe, D., Butler, P., Freeman, S., Greenlees, P. et al. (2003) The GREAT spectrometer. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 204, 634–637.
- Rivest, R.L., Meyer, A.R., Kleitman, D.J., Winklmann, K. & Spencer, J. (1980) Coping with errors in binary search procedures. *Journal of Computer and System Sciences* 20(3), 396–404.
- Saräen, J. (2011) The ion-optical design of the MARA recoil separator and absolute transmission measurements of the RITU gas-filled recoil separator. Ph.D. Thesis, University of Jyväskylä.

**How to cite this article:** Karvanen J, Niilo-Rämä M, Saräen J, Kärkkäinen S. Estimating mean lifetime from partially observed events in nuclear physics. *J R Stat Soc Series C*. 2021;00:1–24. <https://doi.org/10.1111/rssc.12519>

## APPENDIX A

### A.1 | Unbiased but inconsistent estimator

We show that estimator (5), a straightforward extension of estimator (4) for data with partially observed departures, is unbiased but not consistent. The estimator (5) can be written in the form

$$\frac{1}{n} \left( \sum_{i=1}^n 2R_i d_i - \sum_{i=1}^n a_i \right), \quad (\text{A1})$$

where the constant 2 is needed to compensate for the unobserved departures and  $R_i$  is a missingness indicator for which  $P(R_i = 1) = P(R_i = 0) = 0.5$ . The data are missing completely at random which means that random variable  $R_i$  is independent from all other variables.

The estimator (A1) can be written in the form

$$\begin{aligned} \frac{1}{n} \left( \sum_{i=1}^n 2R_i d_i - \sum_{i=1}^n a_i \right) &= \frac{1}{n} \left( \sum_{i=1}^n 2R_i (a_i + x_i) - \sum_{i=1}^n a_i \right) \\ &= \frac{1}{n} \left( \sum_{i=1}^n (2R_i - 1) a_i + \sum_{i=1}^n 2R_i x_i \right). \end{aligned} \quad (\text{A2})$$

The estimator is unbiased for  $\mu$  because the independence of  $R_i$  from  $a_i$  and  $x_i$  implies that

$$\text{E} \left[ \sum_{i=1}^n (2R_i - 1) a_i \right] = 0$$

and

$$\text{E} \left[ \sum_{i=1}^n 2R_i x_i \right] = n\mu.$$

To see that the estimator is not consistent, we consider the variance of estimator (A1) starting from the form Equation (A2)

$$\begin{aligned} \text{Var} &\left[ \frac{1}{n} \sum_{i=1}^n (2R_i - 1) a_i + \frac{1}{n} \sum_{i=1}^n 2R_i x_i \right] \\ &= \text{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n (2R_i - 1) a_i + \frac{1}{n} \sum_{i=1}^n 2R_i x_i - \mu \right)^2 \right] \\ &= \text{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n (2R_i - 1) a_i \right)^2 \right] + 2\text{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n (2R_i - 1) a_i \right) \left( \frac{1}{n} \sum_{i=1}^n 2R_i x_i - \mu \right) \right] \\ &\quad + \text{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n 2R_i x_i - \mu \right)^2 \right]. \end{aligned}$$

Now the first and the third expectations are positive and the second expectation converges in probability to 0 by Slutsky's theorem when  $n$  approaches infinity. We consider therefore only the first expectation that simplifies to

$$\begin{aligned} &\text{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n (2R_i - 1) a_i \right)^2 \right] \\ &= \frac{1}{n^2} \sum_{i=1}^n \text{E} [(2R_i - 1)^2 a_i^2] + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=i+1}^n \text{E} [2(2R_i - 1) a_i (2R_j - 1) a_j] \\ &= \frac{1}{n^2} \sum_{i=1}^n \text{E} [(2R_i - 1)^2] \text{E} [a_i^2] = \frac{1}{n^2} \sum_{i=1}^n \text{E} [a_i^2] \end{aligned}$$



because the independence of  $R_i$  and  $R_j$ ,  $i \neq j$ , implies  $E[(2R_i - 1)(2R_j - 1)] = 0$  and the independence of  $a_i$  and  $R_i$  allows us to factorize the first term. If the sequence  $a_1, a_2, \dots, a_n$  increases linearly on average, the variance increases as a function of  $n$ . As a demonstration, let the time between subsequent arrivals to be a constant  $\mu_a$  so that we have  $a_i = \mu_a i$ . Now the expression increases linearly as a function of  $n$

$$\frac{1}{n^2} \sum_{i=1}^n a_i^2 = \mu_a^2 \sum_{i=1}^n i^2 = \frac{\mu_a^2}{n^2} \frac{n(n+1)(2n+1)}{6} = O(n).$$

We conclude that estimator (A1) is not consistent because its variance increases as a function of  $n$ . In practice this means that estimator (A1) cannot be used. Instead, Section 4 describes an approach where the complete data estimator (4) is applied with a bias correction.

## A.2 | Expected number of departures after the activity period

We derive the formula (8) for the expected number of post-activity departures. Let  $t$  be the activity time and assume that the follow-up continues without limits. The arrivals follow a Poisson process with rate  $\lambda_a$  and the lifetimes are exponentially distributed with mean  $\mu$ . For an arrival at time  $a$ , the corresponding departure occurs after  $t$  with probability  $P(d > t|a) = \exp(-(t-a)/\mu)$ . As the arrivals are uniformly distributed on the interval  $[0, t]$  and independent from the lifetimes, we obtain the unconditional probability

$$P(d > t) = \int_0^t e^{-(t-a)/\mu} \frac{1}{t} da = \frac{\mu - e^{-t/\mu}}{t}. \quad (\text{A3})$$

Formula (8) is obtained when Equation (A3) is multiplied by  $\lambda_a t$ , the expected number of arrivals, and it is taken into account that only half of the departures are observed.