Nuclear Inst. and Methods in Physics Research, A 889 (2018) 16–21

Contents lists available at ScienceDirect

Nuclear Inst. and Methods in Physics Research, A

journal homepage: www.elsevier.com/locate/nima

Matching problem for primary and secondary signals in dual-phase TPC detectors

B. Radics [a,*], E. Burjons [b], A. Rubbia [a]

[a] ETH Zürich, Institute for Particle Physics and Astrophysics, CH-8093 Zürich, Switzerland
[b] ETH Zürich, Department of Computer Science, CH-8092 Zürich, Switzerland

ARTICLE INFO

Keywords:
Liquid scintillator
Dual phase
Time projection chamber
Delayed signal
Matching problem

ABSTRACT

The problem of matching primary and secondary light signals, belonging to the same event, is presented in the context of dual-phase time projection chambers. In large scale detectors the secondary light emission could be delayed up to order of milliseconds, which, combined with high signal rates, could make the matching of the signals challenging. A possible approach is offered in the framework of the Stable Marriage and the College Admission problem, for both of which solutions are given by the Gale–Shapley algorithm.

1. Introduction

Time projection chambers (TPC) using liquid scintillator are used for three dimensional event reconstruction [1]. Liquid TPC detectors operate with an electric field across the liquid, and as a result of an ionization by a primary particle, liberated electrons can be drifted up to the order of meters. A variant of the liquid TPC detectors is the dual-phase (DP) TPC, which allows for additional charge amplification of the drifted electrons by employing a strong electric field across a small gas gap above the liquid phase [2]. This realization produces primary and secondary signals, the latter being delayed due to the finite drifting of electrons in the liquid phase. Measuring the time of delay of the secondary signal encodes information on the drift distance, and therefore provides the position of the event. The electron drift velocity, however, may be on the order of $\sim$ mm/μs, leading to a delay of the secondary light signal up to the order of milliseconds for large scale liquid scintillator experiments, which complicates the reconstruction of the events. Because the sensitivity of the measurements scales with the active target size the upcoming DP TPC detectors will have sizes in the order of several meters, which automatically introduces the problem of how the matching of the primary and secondary light signals can be performed for single or multiple scatter events. Experiments that may face such scenario are typically based on large scale liquid scintillator installation such as the DUNE far detector [3,4], DarkSide-20k [5], and ArDM [6].

In this work we provide a possible approach to the DP TPC event matching problem using an algorithm from Gale and Shapley [7], which gives a solution to the Stable Marriage problem. Matching, in mathematical sense, is selecting a set of independent edges in a graph without common vertices [8]. When applying the concept to DP TPC data, vertices are assumed to be the events and edges are the possible combinations of the events. The problem is to find matching between two classes of events (primary and secondary signal events) in such a way that there are no two events from one class, which could both belong to the same event from the other class. For single scatter events, in which a particle deposited energy only once in the detector, this matching scheme can be applied. For multiple scatter events, other algorithms may also be considered as there could be additional secondary events belonging to the same primary event. Gale and Shapley also provided a solution to a similar problem, the so-called College Admission problem [7]. In this paper, we study the application of the Stable Marriage and College Admission algorithms both to single and multiple scatter events using data generated by a toy Monte Carlo.

2. The Stable Marriage and the college admission problems

2.1. The Stable Marriage problem

The problem of matching prompt and delayed secondary signals is similar to that of assigning members of a group to another group based on the individual ranking of the group members. The context of the paper of Gale and Shapley is how to reach stable marriages (only in mathematical sense) between a set of men and women, each having their own ranking for the opposite gender candidates. The algorithm consists of an iterative procedure which stops when an actual stable set of marriages is found. The men first propose to the first ranking women

* Corresponding author.
E-mail address: bradics@phys.ethz.ch (B. Radics).

https://doi.org/10.1016/j.nima.2018.02.022
Received 24 September 2017; Received in revised form 9 December 2017; Accepted 4 February 2018
Available online 7 February 2018

on their list, and each woman rejects all but her favorite from those who proposed to her. The women then keep their selected candidate on a string to allow the possibility of a proposal from higher ranking men on their list in the next iteration. The rejected men will propose in the following iteration to the next candidate on their rankings. The same proposal/rejection routine is repeated iteratively until all women have been proposed to. The algorithm is stable, that is by the end there is no such pair of unmatched man and woman who would prefer each other over their previously established candidates. This algorithm and its adaptations have been successfully applied in the assignment of medical students to universities, job assignments, roommate selections among others [8]. Variations include: unweighted graphs, where there is no preference list for the candidates, assigning multiple pairings to one vertex (college admission, see below) and matching in non bipartite graphs, where there is only one type of data (roommate selections) [9,10]. The criteria for a good solution might also vary. For example, the Hungarian algorithm [11,12] finds the matching with the highest likelihood even if it is not stable. This algorithm and its variations have also been successfully implemented and its performance improved. All of these together with other approaches appear in the books by Gusfield–Irving [9] and Knuth [10].

2.2. The College Admission problem

The use case of matching multiple members of a group to single members of another group is covered by the solution to the College Admission problem. Briefly, students apply to a certain number of colleges, each with certain quotas and rankings for the students, while each student having their own ranking for the colleges. In the first iteration of the solution, students apply to the first colleges on their ranking lists, and the colleges take a number of students according to the available quotas, and put these students on a waiting list, while the rest of the students are rejected. In the next iteration, the previously rejected students apply to the second colleges on their ranking lists. The colleges consider the new applicants and compare them with those on their waiting lists, picking only the top students from the two sets according to their ranking lists, while rejecting the rest. The iteration terminates when every student is on a waiting list or has been rejected from all colleges. At this point all students on the waiting lists are admitted to the colleges.

3. The delayed secondary signal matching problem with single scatter events

The context presented can be translated into the problem of matching primary and delayed secondary signals from the same event. The situation is illustrated in Fig. 1. Events are recorded sequentially in time order and are assumed to be classified to be primary (S1) or secondary (S2) signal type. Upon detecting a secondary signal event a ranking list can be constructed, containing a value for each of a set of previously detected primary signal events such that the first in the ranking list is the most compatible to match the secondary signal to. The ordering rule in the ranking must be based on some information commonly shared by the event characteristics (such as event topology, etc.). The Gale–Shapley algorithm gives a stable solution given any plausible ranking table, but it is not necessarily the true solution (experimental conditions, such as attenuation of drifting charge or inefficient extraction of electrons from the liquid phase, might be present in the data). The ordering rule plays a key role in finding the best matching. In addition to a experimentally validated ordering rule, the primary and delayed secondary signal candidates must also be present in the data set analyzed, which might put some constraints on the size of the data set used to search for the matching signal.

In any case, assuming the above conditions are met, an ordering rule can always be constructed (e.g. as a Likelihood function), which gives a measure of compatibility between signal events. In this way a ranking list can be made for each primary (secondary) signal events with respect to all later (earlier) opposite type events. In the following we use a Likelihood ordering and illustrate a simplified situation of toy Monte Carlo simulated primary and delayed secondary signals.

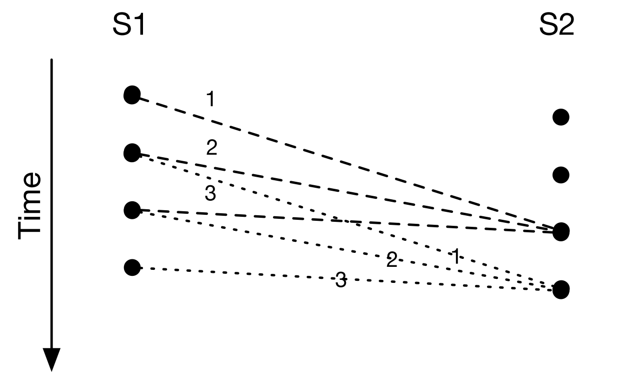

Fig. 1. Illustration of matching possibilities of primary (S1) and secondary (S2) events to each other. The dashed and dotted lines indicate candidate edges for fixed S2 events to multiple S1 candidate events. The numbers on the edges indicate possible ranking orders which is to be estimated from detection specific information.

3.1. Application to single scatter toy Monte Carlo simulation

Here we assume a simplified scenario of a DP TPC with 1.5 m drift length, electron drift velocity $v_{\mathrm{drift}} = 1.3$ mm/μs. The primary signals (S1) are generated randomly and uniformly across the 1.5 m long detector, and for each primary signal, the secondary signal (S2) is generated after a delay time given by the electron drift velocity and the distance of the primary event from the liquid surface. The position of the primary event may be smeared by a Gaussian in order to mimic the detector response. In the following, a Likelihood function is constructed in order to perform the ranking of events for matching. Most experiments can recover some rough event topology information from the relative amount of light detected in various subdetectors. We assume this is the case also in this simplified scenario, which means that a rough guess can be made on the position, x_{s1}, of the primary signal event in the liquid. Therefore having a particular secondary signal event, S2, detected for each previously detected primary signal event, S1, the Likelihood of being the correct match can be calculated formally as,

$$L(S1|S2) \propto g(x_{s1}, t_{s1}|x_{s2}, t_{s2}, v_{\mathrm{drift}}), t_{s1} < t_{s2} \quad (1)$$

In the above formulation, g denotes an arbitrary measure of probability, but for a simplified case a Gaussian is assumed, whereby signal pairs with calculated delay time $dt' = (x_{s2} - x_{s1})/v_{\mathrm{drift}}$ closer to the observed values of $dt = t_{s2} - t_{s1}$ get a higher probability.

As a toy example, 5 events are presented for primary and secondary signals, generated with 1 kHz fictitious rate for the primary signal. Events are shown in order of the time of their detection in Table 1. In the table, x indicates the calculated position for the event from the subdetector information, and for each S2 delayed secondary signal the corresponding true S1 primary event is also indicated. The generated events demonstrate the property that sometimes a primary signal might be detected before the secondary signal arrived for the previous primary event.

The ranking order for these events has been calculated using the Likelihood function given in Eq. (1). The corresponding ranking tables are presented in Tables 2 and 3, respectively. The condition that the delayed secondary signal must happen, by construction, later than the primary signal is explicitly visible in the empty cells in the ranking tables. In the example we gave full ranking tables, however, in practice