



**HAL**  
open science

## Robust label attribution for real-time bidding

Martin Bompaire, Antoine Désir, Benjamin Heymann

► **To cite this version:**

Martin Bompaire, Antoine Désir, Benjamin Heymann. Robust label attribution for real-time bidding. 2021. hal-02998099v3

**HAL Id: hal-02998099**

**<https://hal.science/hal-02998099v3>**

Preprint submitted on 16 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Robust label attribution for real-time bidding

Martin Bompaire

Criteo AI Lab, [m.bompaire@criteo.com](mailto:m.bompaire@criteo.com)

Antoine Désir

Technology and Operations Management, INSEAD, [antoine.desir@insead.edu](mailto:antoine.desir@insead.edu)

Benjamin Heymann

Criteo AI Lab, [b.heyman@criteo.com](mailto:b.heyman@criteo.com)

This version: June 16, 2021

Most of the display advertising inventory is sold through real-time auctions. The participants of these auctions are typically bidders (Google, Criteo, RTB House, Trade Desk for instance) who participate on behalf of advertisers. In order to estimate the value of each display opportunity, they usually train advanced machine learning algorithms using historical data. In the labeled training set, the inputs are vectors of features representing each display opportunity and the labels are the generated rewards. In practice, the rewards are given by the advertiser and are tied to whether or not a particular user converts. Consequently, the rewards are aggregated at the user level and never observed at the display level. A fundamental task that has, to the best of our knowledge, been overlooked is to account for this mismatch and split, or attribute, the rewards at the right granularity level before training a learning algorithm. We call this the label attribution problem.

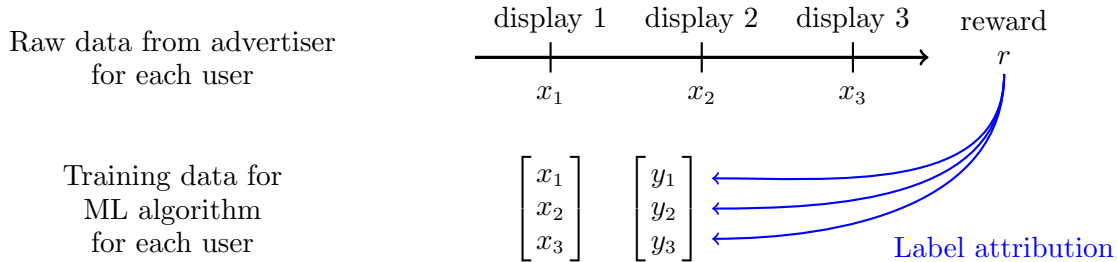
In this paper, we develop an approach to the label attribution problem which is both theoretically justified and practical. We dub our solution the robust label attribution because it satisfies several desirable properties, including distributional robustness. Moreover, we develop a fixed point algorithm that allows for large scale implementation and showcase our solution using a large scale publicly available dataset from Criteo, a large Demand Side Platform.

**Key words.** Online advertising, attribution, distributional robustness

---

## 1. Introduction

Digital advertising has been growing continuously since its inception and sustains a large part of the internet as we know it. It is indeed the main source of revenue for several tech giants (Google and Facebook in particular), and has attracted billions of dollars from advertisers over the last two decades. In 2019, the global digital ad spending was more than 300 billion USD (Enberg 2019). What makes digital marketing so attractive is that marketers can harness the power of data to make informed budget allocation decisions, and show the right banner to the right customer at the right time.



**Figure 1** For each user,  $x_i \in \mathbb{R}^d$  is a vector of features representing the context of the  $i^{\text{th}}$  display opportunity. For each user, the bidder receives from the advertiser a reward  $r \in \mathbb{R}$ . The label attribution problem consists of splitting the reward  $r$  given by the advertiser for a given user into labels  $y_i \in \mathbb{R}$  for each of the display opportunity for that user.

In this paper, we take the point of view of a *bidder* (Google, Criteo, RTB House, Trade Desk for instance) who participates in a real-time bidding exchange on behalf of an advertiser (Balseiro and Candogan 2017). Usually, such a bidder is rewarded when an ad is won, clicked on and followed in a short period of time by a conversion, typically a sale or any kind of action on the advertiser website such as creating an account or putting an item in a basket. Whenever the bidder receives a display opportunity, he has to estimate its potential reward and then, whatever the auction mechanism, calibrate his bid accordingly. The prediction of this potential reward is usually the result of a supervised learning algorithm that requires *labeled* training samples where the inputs are vectors of features representing the display opportunity (these features encode contextual and user information) and the outputs are the generated rewards. Unfortunately, the data does not readily come in this form in practice. Indeed, the rewards, which are decided by the advertiser, are tied to a particular user. More precisely, for each user, the advertiser decides, among the different advertising channels (Demand-side platforms, agencies, search) that the user interacted with, which one gets the credit for the conversion. From the bidder’s perspective, this implies that the labels are given for sequences of inputs and not directly for each individual input (see Figure 1). Existing approaches (Chapelle 2014) omit this important step of converting the raw data given by the advertiser into a dataset that can be fed to a machine learning pipeline and attribute, somewhat heuristically, the reward given by the advertiser to the label of the last clicked display. In this paper, we propose a principled and scalable solution to predict the value of a given display opportunity that accounts for this mismatch between the raw data given by the advertiser and the training data needed for a machine learning (ML) pipeline.

This problem is related to the well studied question of attribution in online advertising, i.e. assigning conversions credits to individual marketing interactions. It is usually posed from the point of view of the advertiser who needs to decide how to split the credit between different advertising channels and is a fundamental questions that informs the media mix optimization or

the understanding of a customer’s journey. It has recently been identified as a top research priority by the Marketing Science Institute (Lemon 2016). Similarly, in deciding how to split the reward given by the advertiser for sequences of display opportunities, the bidder also needs to perform his own attribution. To differentiate between these two levels of attribution, we refer to the bidder’s task as the *label attribution* problem. For our purposes, we assume in this work that the attribution mechanism used by the advertiser, which impacts the reward of the bidder, is fixed and exogenous for the bidder. This leader-follower assumption is mild since the advertiser attribution rule is almost always fixed in practice. In a way, the label attribution is a more granular attribution done by the bidder. The bidder needs to choose a label attribution to maximize a reward given by the advertiser attribution. This makes these two attributions slightly different in spirit. We would like to also remark that even if the advertiser attribution is done at the display level, this is sometimes referred to as multi-touch attribution, this distinction is still useful as it is not clear that this display level attribution chosen by the advertiser is optimal for the bidder who might gain by recomputing his own internal attribution. Among the technical challenges a bidder is facing, choosing a label attribution mechanism is often an understated problem.

### 1.1. Motivating example

We present an example to illustrate the limitations of ignoring the label attribution step and motivate the need for a more principled approach. Consider a bidder who is trying to evaluate different display opportunities. In order to train a machine learning algorithm, the bidder needs to have, for each display opportunity, an input vector of features  $x$  describing contextual information about the display opportunity and a reward encoded in a label  $y$ . However, the reward data coming from the advertiser is aggregated at the user level and hence does not come at the right granularity level. What is implicitly done in practice is to assign the reward to the last opportunity label. We will refer to this method as the *last touch label attribution*. More precisely, if, for a particular user, the bidder’s data consists of a sequence of display opportunities  $x_1, \dots, x_n$  leading to a reward of  $r$ , then the current practice is to create the following training data:  $(x_1, 0), \dots, (x_{n-1}, 0), (x_n, r)$ . A similar last touch heuristic is also prevalent for the advertiser attribution (Ji and Wang 2017, Diemert et al. 2017). Just like the last-touch attribution used by advertisers, this last touch approach for the label attribution fails to recognize the potential effect of early displays moving a customer up the conversion funnel and therefore fails to acknowledge their contribution into generating the reward. This leads to undervaluing early display opportunities and overvaluing later ones.

To make this even more concrete, assume for simplicity that there are only two types of display opportunities. Displays A influence users’ behavior and increase their conversion probability. On the other hand, displays B do not influence the conversion probability at all but artificially always

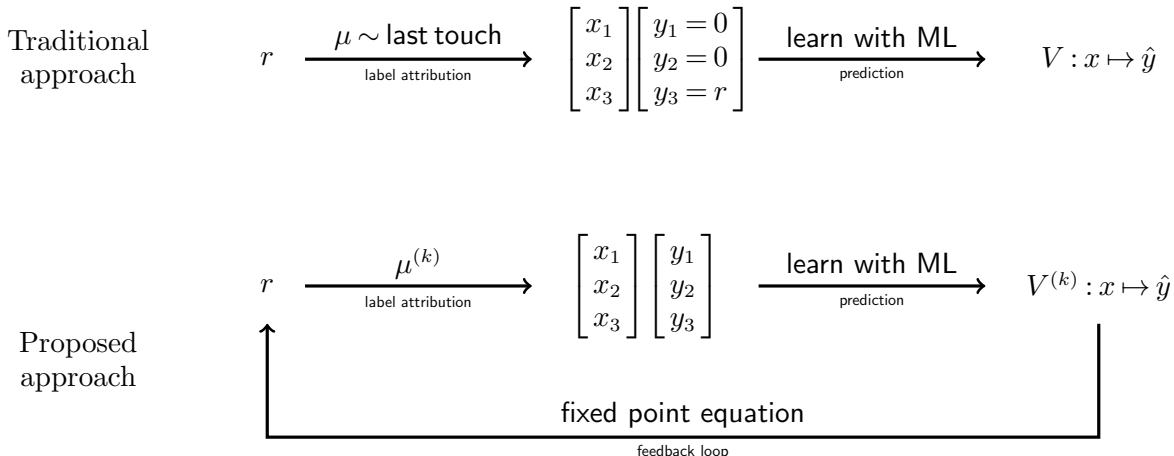
appear before a conversion, as is common for search ads (Blake et al. 2015). In such situation, the last touch label attribution wrongly allocates all the reward to displays B. Over time, the bidder will undervalue displays A and overvalue displays B. This in turn will lead to less conversions and therefore also lower the reward attributed to the bidder. This example highlights that the label attribution is a crucial step to align the bidder strategy with the given reward. It showcases the limitations of the current last touch approach and motivates the development of alternate methods of label attribution. We end this example by noting that the motivation for using a last touch label attribution rule in practice seems to be driven by the misconception that if the advertiser uses a last touch attribution rule, then the bidder should in turn use a last touch label attribution rule. Since our example does not depend on the particular advertiser attribution rule, it demonstrates that this belief is wrong.

## 1.2. Contribution

In order to accurately bid in any auction, a bidder needs to evaluate each display opportunity. The main challenge is that rewards are given at the user level, i.e. for sequence of displays, whereas the contribution of each individual won display is never observed. Consequently, we introduce in this paper the label attribution problem which consists of finding a mapping  $\mu$  which splits the reward for every user into labels for each individual display opportunity. Our first contribution is to identify and formalize this label attribution problem which has not, to the best of our knowledge, been studied before. The exact model and notations are introduced in Section 3.

A natural way to value a display is through its marginal contribution, i.e. the lift it provides to an existing history of displays. Mathematically, we formalize this notion by imposing that the valuation function  $V$ , used to evaluate display opportunities, must be additive. This property defines a unique label attribution mechanism which we dub the *robust label attribution*. We then show that this intuitive label attribution enjoys several interesting properties. First, we show that bidding according to the robust valuation is myopic optimal, i.e. optimal if there are no other future opportunities. Even though this notion may seem rather narrow, it alleviates some limitations of the last touch approach. In particular, we show that it correctly learns the values in the motivating example from Section 1.1. Moreover, we show that our proposed approach is distributionally robust, i.e. does not depend on how many times each history has been observed. This is very useful from a learning perspective and this property justifies the name robust label attribution. Finally, we show that the robust label attribution satisfies a fixed point equation which we leverage to design an efficient learning algorithm.

In an ideal setting where the bidder has data on every possible history, the additivity property is enough to compute the robust label attribution. However, in practice, the data is incomplete.



**Figure 2** The traditional label attribution approach uses the last touch heuristic. By contrast, our proposed approach is an iterative algorithm which alternates between updating the attribution mapping  $\mu^{(k)}$  and the prediction function  $V^{(k)}$ . The update of the mapping  $\mu^{(k)}$  relies on a fixed point relation. Our framework can be implemented by adding a feedback loop to an existing ML pipeline.

We build on the fixed point equation satisfied by our robust label attribution to design an efficient learning algorithm. Our proposed approach overcomes the sparsity of the available data by leveraging the high-dimensionality of the available feature space. Our method is very practical as it can be embedded in an existing ML pipeline. More precisely, our approach is iterative and alternates between an update step and a predict step. In each iteration  $k$ , we maintain two mappings: a label attribution mechanism  $\mu^{(k)}$  which is used to split the rewards, and a valuation mechanism  $V^{(k)}$  which transforms any vector of features into a predicted reward. In the first step of each iteration, we update the mapping  $\mu^{(k)}$  using the fixed point relation and the learned mapping  $V^{(k)}$ . We then, in a second step, split the rewards from the raw data using the label attribution  $\mu^{(k+1)}$  and learn a new mapping  $V^{(k+1)}$  using any ML algorithm. We show that the sequence of mapping  $\mu^{(k)}$  converges to our robust label attribution. Figure 2 illustrates our framework.

Finally, we illustrate the practicality of our robust label attribution framework on a large scale dataset<sup>1</sup>. Our method outperforms the last touch approach on various metrics. Moreover, we show that the proposed attribution is able to capture some key qualitative relations of the problem. For instance, in the learned values, the ads exhibit a diminishing marginal return which is a phenomenon that has been observed in practice.

**Agenda.** The paper is organized as follows. We begin with a review of the literature in Section 2. We then introduce the notations that allow us to properly frame the label attribution problem in Section 3. In Section 4, we present our main methodological contribution consisting of several

<sup>1</sup> All our codes are available at <https://github.com/criteo-research/robust-label-attribution>

characterizations of the robust label attribution together with an explicit algorithmic framework. Finally, we illustrate our ideas on a large publicly available dataset in Section 5.

## 2. Related Work

We survey the immediate related literature and refer to Choi et al. (2020) and Wang et al. (2017) for more background on display advertising.

Most of the advertising is done through auctions, and we thus take the point of view of a bidder in this paper. Regardless of the auction mechanism, the bidder has to estimate the value he could generate from each display opportunity and then calibrate his bid accordingly. For example, in a second price setting, for a given display opportunity, the standard in the industry is to bid the estimated average value earned with this specific impression (Perlich et al. 2012, Wang et al. 2017). This is referred to as *value based bidding* in Xu et al. (2016). We would like to emphasize that this behavior is *myopically* designed in the sense that the future policy is disregarded or equivalently that the bid is computed as if this was the last opportunity to generate a conversion. Our robust label attribution adopts a similar philosophy and offers a principled myopic optimal approach. Without loss of generality, we assume that the bidder is compensated in proportion to the reward, i.e. the advertiser attribution, he received. Note that the bidder might need to compute some shading factor to adapt to the competition and the auction mechanism (Balseiro and Gur 2019, Balseiro et al. 2021). However, the computation of such shading is beyond the scope of this paper.

Realizing the need for more principled solutions to the advertiser attribution problem, recent works have started proposing algorithmic approaches to go beyond rule-based heuristics such as last touch. Using different generative models, Danaher (2018) and Anderl et al. (2016) both propose a counterfactual approach. The notion of Shapley value (Berman 2018, Singal et al. 2019, Dalessandro et al. 2012) has also been used to split the credit fairly among different channels. Johnson et al. (2017), Lewis et al. (2011), Lewis and Rao (2015) all study models related to causality and advertiser attribution. Their approach is based on econometric parametric models.

Our research is inspired by the multi-touch attribution (MTA) literature (Shao and Li 2011, Ji et al. 2016, Arava et al. 2018). This stream of work aims at designing machine learning algorithms that assign fractional credits to different touchpoints from the advertiser’s point of view. They also aim at solving the advertiser attribution problem but focus on the implementability of the solution rather than its theoretical justification. Despite using a different formalism, our work relates to the survival analysis approach used in Laub et al. (2015). Moreover, the fixed-point-based algorithm of Section 4.2 shares some characteristics with the EM algorithm in Zhang et al. (2014). Similarly, Ji and Wang (2017) relies on a Weibull model and is very close to our work. However, their choice of parametrization hides the fixed-point property identified in Zhang et al. (2014).

On top of the advertiser attribution problem, some researchers have also considered the bidder’s problem as we do and tried to improve the buying process using machine learning pipelines. Chapelle et al. (2014) presents a bidding architecture which relies on an implicit last touch label attribution. Chapelle (2014) provides a technique to deal with the delay between displays and sales. Diemert et al. (2017) proposes to model the advertiser attribution to increase bidding performances. Their experimental results motivate our approach. More recently Bompaire et al. (2020) provides a framework based on reinforcement learning and causal inference. Their approach – heuristic and forward-looking – is complementary to ours – axiomatic and distributionally robust.

Finally, we would like to highlight that the question of attribution is not specific to online advertising. For instance, recent work in interpretable machine learning tries to attribute the prediction of a machine learning model to individual features of the input (Dhamdhere et al. 2018). In a very different context, Flores-Szwagrzak and Treibich (2020) develops a method to disentangle individual performance from team productivity. In each of these examples, the goal is to attribute some aggregate output to individual components of the input. Although Flores-Szwagrzak and Treibich (2020) develops an iterative algorithm which is similar in spirit to our approach, their problem of splitting credit among team members is static in nature, while time is an important aspect of our problem. Another important difference is that we propose a machine learning implementation of our algorithm (see Section 5) which is able to leverage the presence of context features.

### 3. Framework

The setting is as follows (Golrezaei et al. 2018, Heymann 2019, Choi et al. 2020).

1. *Learning step*. The bidder receives a set of user histories, where each history is associated with a reward, typically the result of the advertiser attribution.
2. *Bidding step (online)*. The bidder receives bid requests and, for each of them, needs to estimate from the context features the display opportunity value (or *valuation*), i.e. the potential contribution toward the reward. The bidder then uses this valuation to compute the final bid.
3. *Reward step*. The bidder is rewarded proportionally to the advertiser attribution it receives. More precisely, note that in order for a user history to generate a reward, the user needs to convert and the advertiser needs to attribute this conversion to the bidder.

As mentioned in the introduction, the challenge is to compute a valuation for a display opportunity, given that the rewards are given at the user timeline level. This section aims at formalizing the label attribution problem depicted in Figure 1. Even though this problem naturally appears in an engineering settings where one needs to construct a labeled training set, it has not, to the best of our knowledge, been identified as a research topic. Our formalism shed some clarity on this topic which we believe is often misunderstood in the industry.



### 3.1. Model and notation

We assume that each display is characterized by a vector of features  $x \in \mathbb{R}^d$  and denote by  $\mathbb{D} \subseteq \mathbb{R}^d$  the set of past and future displays opportunities. The features correspond to different products or creatives as well as encode contextual information such as whether or not the user has already seen an ad. In general, any feature that is useful for the bidder can be encoded in the display set  $\mathbb{D}$ . A **history**  $h$  is a nonempty sequence of elements of  $\mathbb{D}$  that represents a sequence of displays – with their characteristics – shown to a user. For any history  $h$ , we denote by  $|h|$  its size and use the notation  $h = [h_1, \dots, h_{|h|}]$ , where  $h_i \in \mathbb{D}$  for  $i \in \{1, \dots, |h|\}$ , to denote its individual elements. Furthermore, we write  $h \preceq h'$  if  $h$  is a sub-history of  $h'$ , that is if  $h$  is equal to the sequence of the  $|h|$  first displays of  $h'$ . Let  $\mathbb{H}$  be the set of all possible histories over which we are given a probability  $\mathbb{P}$  that denotes the propensity of each history to appear in the historical logs. Note that our notion of history does not depend on any underlying generative model: The distribution  $\mathbb{P}$  can directly be computed from data or can come from a particular generative model. This makes our approach quite robust.

Given our focus on display advertising, we assume that  $\mathbb{D}$  represents a set of displays. However, it could in principle represent any type of marketing interactions. We also present in Appendix E an extension of the model which can account for additional events that cannot be captured in the feature space such as whether or not a display is going to be clicked on. For any history  $h$ , we let  $h_{[1, \dots, i]}$  be the projection of  $h$  onto its first  $i$  elements, i.e. for all  $i \leq |h|$ , we have  $h_{[1, \dots, i]} = [h_1, \dots, h_i]$ . With these notations,  $h' \succcurlyeq h$  is equivalent to  $|h'| \geq |h|$  and  $h'_{[1, \dots, |h|]} = h$ . Finally, for all  $x \in \mathbb{D}$  and  $h \in \mathbb{H}$ , let  $h + [x] = [h_1, \dots, h_{|h|}, x]$  denote the history obtained from adding a display  $x$  to a history  $h$ . We end this section with two technical assumptions.

ASSUMPTION 1.  $\mathbb{D}$  is a finite set.

Assumption 1 is not limiting in practice. Indeed, if some display features are continuous rather than categorical, they can be quantized (Chapelle et al. 2014) in order to satisfy Assumption 1. We also make the following assumption on the underlying probability distribution of histories.

ASSUMPTION 2. If  $\mathbb{P}$  has support on a history, then it also has support on all its sub-histories. Namely, if  $\mathbb{P}(h) > 0$  for some  $h \in \mathbb{H}$ , then  $\mathbb{P}(h') > 0$  for all  $h' \preceq h$ .

This is a mild assumption which is verified for various generative models used in the literature such as Markovian models (Archak et al. 2010, Anderl et al. 2016, Singal et al. 2019) or point process models (Xu et al. 2014).

### 3.2. Reward

We associate each history with a *reward*, corresponding to its value. From the bidder perspective, this value comes from a black-box decided by the advertiser. In the case where the bidder and advertiser are the same, these rewards could represent sales or whatever the advertiser is optimizing for. However, in general, the bidder and advertiser are different actors.

**DEFINITION 1 (REWARD).** The **reward function**  $R$  is a mapping from  $\mathbb{H}$  to  $\mathbb{R}^+$  that associates to each history a reward. More precisely, for a history  $h \in \mathbb{H}$ ,  $R(h)$  represents the average reward given by the advertiser to the bidder when he shows the sequences of displays in  $h$ .

In the example from Figure 1, the reward correspond to the value  $r$ . As discussed in the introduction, the reward given by the advertiser might be given at the display level. In this case, the reward of a history should be thought of as an aggregate value, i.e. the sum of values over displays in the history. As argued, it is indeed unclear that the split over displays given by the advertiser is optimal for the bidder. What we advocate here is that the bidder performs his own label attribution on the aggregate value of a history which motivates having one reward for each history. We make the following monotonicity assumption for the reward.

**ASSUMPTION 3 (Non decreasing rewards).**  $\forall h' \succcurlyeq h, R(h') \geq R(h)$ .

This property implies that an additional display never decreases the reward of a history. Importantly, this does not exclude a potential diminishing marginal effect for displaying more ads. In Sections 3 and 4, we assume that we have access to  $R$ . In practice, the function  $R$  is not known and we illustrate in Section 5 how it can be estimated from data.

### 3.3. Valuation

The bidder needs to estimate the value for each new display opportunity. The value assigned to this display should intrinsically be a function of the history. Indeed, the marginal effect of an ad on a customer who is not aware of a product is perhaps more drastic than for a customer who has already seen multiple ads for this product.

**DEFINITION 2 (VALUATION).** A **valuation**  $V$  is a mapping from  $\mathbb{D} \times \mathbb{H}$  to  $\mathbb{R}^+$  used by the bidder to characterize the value of a display given a particular history. More precisely, for  $(x, h) \in \mathbb{D} \times \mathbb{H}$ , we denote by  $V(x|h)$  the expected value of the display  $x$  given the history  $h$ .

Using this notation, if a user's history is  $h = [x_1, x_2]$  for some  $(x_1, x_2) \in \mathbb{D}^2$  as in the example introduced in Figure 1, the bidder uses  $V(x_3|h)$  to decide how much he should be ready to pay for the display  $x_3 \in \mathbb{D}$ . The bidder's challenge is to appropriately compute a valuation  $V$  from the reward function  $R$ . A key intermediate step is the label attribution which we discuss next since it allows to create the training set needed to learn  $V$  as the result of a ML algorithm.

### 3.4. Label attribution

We define the label attribution as the process of dividing the reward of a particular history among its individual displays. This is a natural step as the rewards are given for entire histories but any bidding strategy has to evaluate adding a particular display to an ongoing history. Using a ML terminology, the reward corresponds to a label at the history level. For bidding purposes, the bidder needs to predict the value of individual displays, and therefore convert those rewards at the history level into labels at the display level in order to train a ML model. This is precisely what the label attribution does.

**DEFINITION 3 (LABEL ATTRIBUTION).** A **label attribution**  $\mu$  is a mapping from  $\mathbb{N} \times \mathbb{H}$  to  $\mathbb{R}^+$  that satisfies the following properties.

1.  $\forall h \in \mathbb{H}, \forall i > |h|, \mu(i, h) = 0.$
2.  $\forall h \in \mathbb{H}, R(h) = \sum_{i=1}^{|h|} \mu(i, h).$

For a label attribution  $\mu$  and history  $h \in \mathbb{H}$ ,  $\mu(i, h)$  represents the label of the  $i^{\text{th}}$  display  $h_i$ , i.e. the part of the reward  $R(h)$  attributed to  $h_i$ . The first property states that we can only split the reward of a history among its displays. The second property states that the attribution has to be balanced and does not attribute more than the reward  $R(h)$ . As a concrete example, with our notation, the last touch label attribution can be written for all  $(i, h) \in \mathbb{N} \times \mathbb{H}$  as

$$\mu^{\text{LT}}(i, h) = \begin{cases} R(h), & \text{if } i = |h|, \\ 0, & \text{otherwise.} \end{cases}$$

Just like the advertiser's attribution, the label attribution is splitting credit. However, the split is done at the more granular level of the displays. Indeed, the bidder needs to predict the value of a display opportunity on the fly before the end of the history. Going back to the bidder's strategy, each label attribution is naturally associated with a valuation as we discuss next.

### 3.5. Associated valuation

Every label attribution  $\mu$  naturally induces a valuation. Indeed, for a given history  $h$ , recall that the valuation  $V(x|h)$  aims at capturing the value of the display  $x$  when added to the history  $h$ . However, the bidder does not know whether additional opportunities will come in the future. If no other opportunity arises, then the display should be valued  $\mu(|h|, h)$ . On the other hand, if more displays are added, then the display should be valued  $\mu(|h|, H)$  for some  $H \succ h$ . It is natural to introduce the valuation associated with the label attribution  $\mu$  obtained by averaging on the future possibilities.

**DEFINITION 4 (ASSOCIATED VALUATION).** We say that a valuation is associated to the label attribution  $\mu$ , and denote it by  $V_{\mathbb{P}}^{\mu}$ , if for all  $(x, h) \in \mathbb{D} \times \mathbb{H}$ , we have

$$V_{\mathbb{P}}^{\mu}(x|h) = \mathbb{E}_{H \sim \mathbb{P}} [\mu(|h| + 1, H) \mid H \succ h + [x]]. \quad (1)$$

We remind the reader that  $h + [x] \in \mathbb{H}$  is the history obtained by appending the display  $x$  to the history  $h$  and that  $H \succcurlyeq h + [x]$  means that we are conditioning on all the histories  $H$  starting with the sequence of displays of  $h + [x]$ . For instance, the valuation associated with the last touch label attribution is given by

$$V^{\text{LT}}(x|h) \triangleq V_{\mathbb{P}}^{\mu^{\text{LT}}}(x|h) = R(h + [x]) \cdot \mathbb{P}(H = h + [x] \mid H \succcurlyeq h + [x]).$$

If the label attribution is driven by the last touch philosophy, then an opportunity is evaluated through its probability of exactly preceding a conversion, which is precisely what the quantity  $\mathbb{P}(H = h + [x] \mid H \succcurlyeq h + [x])$  is capturing. It is worth noticing that this valuation depends on the distribution of future displays. This is precisely the caveat observed in the Introduction as this dependence might lead the bidder to overvalue displays that come late in the conversion funnel.

#### 4. The Robust Label Attribution

A natural way to value a display is through its marginal contribution, i.e. the lift it provides to an existing history. We formalize this idea through the notion of additivity which we define next.

DEFINITION 5 (ADDITIVITY). Let  $V$  be a valuation. We say that  $V$  satisfies the additivity property if for all  $(x, h) \in \mathbb{D} \times \mathbb{H}$ ,

$$V(x|h) = R(h + [x]) - R(h).$$

Note that given a reward function, this property uniquely defines a valuation. In the rest of the paper, we refer to the valuation that satisfies the additivity property as the **robust valuation**. This very natural way of evaluating the next opportunity has recently received attention in the industry (Perlich et al. 2012, Wang et al. 2017). It is also sometimes referred to as lift-based valuation (Xu et al. 2016). Our notion of marginality is temporal, i.e. we are quantifying the marginal contribution with respect to the previous actions. This temporal marginality approach is in line with several works on the additivity effect of ads (Shao and Li 2011, Ji and Wang 2017, Zhang et al. 2014, Dalessandro et al. 2012). This is different than a counterfactual marginality, used for instance in Danaher (2018), Anderl et al. (2016), Bompaire et al. (2020), which quantifies the marginal contribution with respect to an alternative action. The counterfactual marginality depends on the policy used in the future, while the temporal marginality does not. Despite its simplicity, we show in this section, through a series of *equivalent characterizations*, that the robust valuation enjoys several interesting and non trivial properties. Before doing so, we show how it leads to a reasonable valuation for the motivating example presented in Section 1.1. In particular, the robust label valuation is different than the last touch valuation.

**Revisiting the motivating example from Section 1.1** Recall that in this example, we have two types of displays A and B where displays of type B do not influence users’ behavior at all. We discussed in Section 1.1 the reasons the last touch method fails to provide a reasonable bid. On the other hand, the robust valuation assigns a valuation of zero to the displays of type B and captures the fact that displays of type B have no effect on users’ behavior. Indeed, these bad displays do not influence conversion probability meaning that  $R(h) = R(h + [x_B])$  (where  $x_B$  stands for a display of type B) which implies that  $V(x_B|h) = 0$  by the additivity property given in Definition 5.

#### 4.1. Equivalent characterizations

In this section, we show that the robust valuation has several notable properties. We begin by proving that bidding according to the robust valuation is the only myopic optimal bidding strategy. We then show that the robust valuation is the only valuation that does not depend on the probability distribution  $\mathbb{P}$  making it very robust from a learning perspective. Finally, we characterize the label attribution associated with the robust valuation through a fixed-point equation.

**4.1.1. Myopic optimality.** We first show that bidding according to the robust valuation is myopic optimal, i.e. optimal if the sequence of displays terminates without any future opportunities.

*PROPOSITION 1.  $V$  satisfies the additivity property **if and only if** it is myopic optimal, i.e. if in a second-price setting, bidding  $V$  is weakly dominant for any competition profile assuming there will be no other future opportunities.*

We defer the proof of Proposition 1 to Appendix A. In practice, heuristics like last touch are partly used because of the belief that they are in some way myopic optimal. We aim to clarify this and derive a principled myopic optimal valuation which, aligned with our previous examples, is going to be different than last touch.

We acknowledge that this notion of optimality is – as the name suggests – rather narrow: the potential combined effects of the decision at stake (the display to be bought) and the future ones are neglected. Nevertheless, we believe this notion has some merit. In particular, the standard notion of optimality implies that the best decision at a given time depends on what will be done in the future. However, solving the complete optimization problem is in practice often intractable and therefore, one can only hope for approximate solutions. By contrast, as we will discuss next, our proposed strategy does not depend on the probability distribution  $\mathbb{P}$  (and hence, on the impact of our bidding strategy in the future), making it more robust. Finally, our aim is also to shift the discussion to principled approaches to label attribution, very much in the same spirit than Singal et al. (2019). Here, we prove that there is a unique valuation which is myopic optimal.

**4.1.2. Distributional robustness.** We next show that the robust valuation does not depend on the underlying probability distribution  $\mathbb{P}$ , i.e. the propensity of each history to appear in the historical logs.

PROPOSITION 2. *V satisfies the additivity property **if and only if** V is an associated valuation which does not depend on  $\mathbb{P}$  but only on the reward function R, i.e., there exists a label attribution  $\mu$  such that  $V = V_{\mathbb{P}}^{\mu} = V_{\mathbb{Q}}^{\mu}$  for any other probability distribution  $\mathbb{Q}$  over  $\mathbb{H}$  that has the same support as  $\mathbb{P}$ .*

We defer the proof of Proposition 2 to Appendix B. Surprisingly, the additivity property also leads to distributional robustness. From a learning perspective, this means that this approach alleviates the need of learning the underlying probability distribution  $\mathbb{P}$ . In particular, if the environment changes and we receive a new probability distribution on the histories  $\mathbb{Q}$ , then the expected reward can still be computed using the formula

$$\mathbb{E}_{H \sim \mathbb{Q}} R(H) = \mathbb{E}_{H \sim \mathbb{Q}} \sum_{i=1}^{|H|} V_{\mathbb{P}}^{\mu}(H_i \mid H_{[1, \dots, i-1]}),$$

even when  $\mathbb{P} \neq \mathbb{Q}$ . We emphasize that this is not true for any other valuation mechanism. Finally, distributional invariance is often related to causality (Peters et al. 2016, Arjovsky et al. 2019), and it would be an interesting research direction to further explore these connections.

**4.1.3. Fixed-point characterization.** We now characterize the label attribution associated with the robust valuation.

PROPOSITION 3. *V satisfies the additivity property **if and only if** V is associated with the label attribution  $\mu$  defined for all  $h \in \mathbb{H}$  such that  $R(h) > 0$  and  $i \in \{1, \dots, |h|\}$  as*

$$\mu(i, h) = \frac{V(h_i \mid h_{[1, \dots, i-1]})}{\sum_{j=1}^{|h|} V(h_j \mid h_{[1, \dots, j-1]})} \cdot R(h). \quad (2)$$

We defer the proof of Proposition 3 to Appendix C. We refer to the label attribution defined by Equation (2) as the **robust label attribution**. We believe that a novelty of our work lies in formulating this fixed-point characterization and that this relation has very practical implications. Indeed, from a practical perspective, the label attribution is the necessary step to create the training datasets needed for any ML pipeline. In particular, in the next section, we leverage the fixed-point characterization to develop an efficient learning procedure which can be implemented at the scale required for display advertising. We would like to note that slight variants of Equation (2) have appeared in the MTA literature. In particular, it comes up as a first-order condition of a model based on survival analysis in Zhang et al. (2014).

If all possible histories are observed with their reward, the valuation  $V$  can be computed by induction using the additivity property from Definition 5. However, building on the fixed-point characterization given in Proposition 3, it is natural to consider the following iterative procedure to compute  $V$  instead:

$$V^{(k+1)}(x|h) = V_{\mathbb{P}}^{\mu^{(k)}}(x|h), \quad \forall (x, h) \in \mathbb{D} \times \mathbb{H}, \quad (3)$$

where for all  $h$  and  $i \leq |h|$ ,

$$\mu^{(k)}(i, h) = \frac{V^{(k)}(h_i | h_{[1, \dots, i-1]})}{\sum_{j=1}^{|h|} V^{(k)}(h_j | h_{[1, \dots, j-1]})} \cdot R(h). \quad (4)$$

Recall that for any  $\mu$  and  $h \in \mathbb{H}$ , the valuation  $V_{\mathbb{P}}^{\mu}(h)$ , defined in Equation (1), is the expected reward, where the expectation is conditioned on  $H \succcurlyeq h$ . We also initialize the procedure with some label attribution  $\mu^{(0)}$ .

**PROPOSITION 4.** *For any  $\mu^{(0)}$  such that  $V_{\mathbb{P}}^{\mu^{(0)}}(x|h) > 0$  for all  $(x, h) \in \mathbb{D} \times \mathbb{H}$ , the procedure defined in Equations (3)-(4) converges. Moreover,  $\mu^{(k)}$  converges to the robust label attribution.*

We defer the proof of Proposition 4 to Appendix D. Note that the technical assumption  $V_{\mathbb{P}}^{\mu^{(0)}}(x|h) > 0$  on the initial valuation is not required to prove convergence but only to prove that the iterates converge to the robust label attribution. Indeed, without this assumption, some coordinates can get stuck at their initial value.

## 4.2. Algorithm

In practice, we do not know the reward function  $R$  but are working with *realized* histories  $h^1, \dots, h^T$  and their associated rewards  $r^1, \dots, r^T$ . Leveraging Proposition 4, we propose a practical procedure that converges to the robust label attribution. The proposed algorithm is iterative and maintains, in each iteration  $k$ , a label attribution  $\mu^{(k)}$  and a valuation  $V^{(k)}$ . Each iteration  $k$  consists of three steps.

1. **Attribution step.** We begin by splitting the  $T$  histories  $h^t$  and rewards  $r^t$  to create a new dataset  $\mathcal{D}^{(k+1)}$  at the display level. More precisely, for each history  $h^t$  and  $i \in \{1, \dots, |h^t|\}$ , we create a vector of features  $x$  for each sub-history  $h_{[1, \dots, i]}^t$ , corresponding to each display opportunity. For each of these vector  $x$ , we create a label  $y$  by splitting the reward  $r^t$  to each sub-histories using the label attribution  $\mu^{(k)}$ . While rarely formalized, this step is traditionally done implicitly using the last touch label attribution (see Figure 2).

2. **Prediction step.** The second step consists in computing an updated valuation  $V^{(k+1)}$ . By definition,  $V$  is the expected value of the label  $y$ :

$$V^{(k+1)} = (h_i | h_{[1, \dots, i-1]}) \mapsto \mathbb{E}_{(x, y) \in \mathcal{D}^{(k+1)}} [y | x = h_{[1, \dots, i]}].$$

**ALGORITHM 1:** Robust Label Attribution Fixed-point Algorithm**Input:** A dataset of histories  $h^t$  and associated rewards  $r^t$  for  $t \in \{1, \dots, T\}$  and an initial attribution

$$\mu^{(0)}$$

**Output:** Robust label attribution  $\mu$  and its associated valuation  $V$ **for**  $k = 0, 1, \dots$  **do**

1. *Attribution step.* Generate dataset by splitting histories into displays and attributing rewards at the display level with the label attribution  $\mu$

$$\mathcal{D}^{(k+1)} = \{(x, y) = (h_{[1, \dots, i]}^t, \mu^{(k)}(i, h^t)) \text{ for } t \in \{1, \dots, T\} \text{ and } i \in \{1, \dots, |h^t|\}\}.$$

2. *Prediction step.* Update the valuation  $V^{(k+1)}$  with one of these two approaches.

(a) *Averaging.*

$$V^{(k+1)} = (h_i | h_{[1, \dots, i-1]}) \mapsto \sum_{(x, y) \in \mathcal{D}^{(k+1)}} y \cdot \mathbf{1}\{x = h_{[1, \dots, i]}\}.$$

(b) *ML algorithm.* Learn to predict the display label  $y$  from the display vector of features  $x$  by training any ML algorithm on  $\mathcal{D}^{(k+1)}$ . Update  $V^{(k+1)}$  to the resulting mapping.

3. *Fixed-point update.* Update the label attribution  $\mu^{(k+1)}$  with the fixed-point characterization

$$\mu^{(k+1)} = (i, h^t) \mapsto \frac{V(h_i^t | h_{[1, \dots, i-1]}^t)}{\sum_{j=1}^{|h^t|} V(h_j^t | h_{[1, \dots, j-1]}^t)} \cdot r^t.$$

**end**

Consequently, in order to predict  $y$  from some vector of features  $x$ , one only needs to average the labels corresponding to the vector of features  $x$  over elements in  $\mathcal{D}^{(k+1)}$ . However, note that this does not allow to predict the label whose vector of features, or history, is not present in the data. This is also impractical when the data is high-dimensional. For that reason, this prediction step is typically done with a ML algorithm such as a logistic regression. More precisely, a ML algorithm is trained to predict the display labels  $y$  from the corresponding vectors of features  $x$  and the resulting mapping constitutes the new valuation  $V^{(k+1)}$ . In Section 5.2, we illustrate the simple averaging approach on a stylized example and the machine learning implementation on a real dataset.

**3. Fixed-point update.** The last step exploits the fixed-point characterization from Proposition 3 to update the label attribution  $\mu^{(k+1)}$  using the updated valuation  $V^{(k+1)}$ .

Algorithm 1 summarizes the procedure.

If the prediction step is done using the averaging approach (Step 2-a), then the procedure converges as stated in Proposition 4. Indeed, even though Proposition 4 assumes access to the true value  $R(h)$ , i.e., the average reward over multiple realizations of  $h$ , the convergence still holds since each realization  $r^t$  is an unbiased estimator of the reward and there is a linear relation between  $R$  and  $V$ , as implied by Definition 5 and Equation (1). Even though we do not have any convergence



guarantee when implementing Step 2-b, we show on a real dataset that Algorithm 1 still converges when implementing Step 2-b in Section 5.3.

Algorithm 1 shows that not only our robust attribution enjoys theoretical characterizations, it is also useful from a practical point of view. Indeed, even when the data is high-dimensional or when Assumptions 1 and 2 are not satisfied, thus prohibiting us from implementing Step 2-a, we can still run Algorithm 1 using Step 2-b. Importantly, using a ML algorithm also allows leveraging any available features. Note that this is anyway what is done in practice where the ML algorithm is often a logistic regression (Chapelle 2014).

Compared to the existing approaches, the main novelties of our approach are (1) explicitly doing the label attribution (Step 1) instead of implicitly using a last touch label attribution, and (2) adding a fixed-point update (Step 3). As also illustrated in Figure 2 in the Introduction, our approach suggests a change in paradigm whereby instead of estimating the valuation from a fixed internal procedure, such as last touch attribution, the optimal pair of valuation and label attribution should rather be iteratively and jointly updated until convergence. Moreover, this feedback loop is a simple change that can easily be implemented in an existing ML pipeline.

## 5. Numerical Experiments

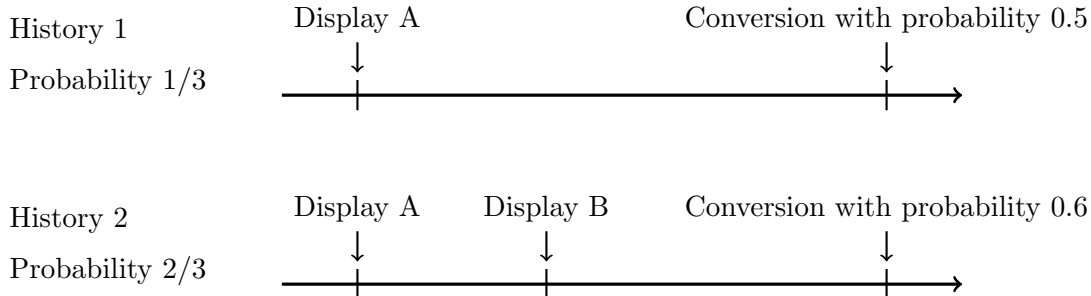
This section presents numerical experiments to test the robust label attribution framework. We first validate the approach on a synthetic dataset in Section 5.2, and then apply it on a large scale publicly available dataset (see Diemert et al. 2017) from Criteo – a large Demand Side Platform – in Section 5.3. In Section 5.1, we begin by defining a likelihood function that measure how “additive” a valuation is. We use this likelihood as a metric to evaluate different valuations or to measure convergence when we use Step 2-b in Algorithm 1 by a ML block.

### 5.1. A likelihood for additivity

The convergence of Algorithm 1 relies on showing that it corresponds to the iterative procedure of a majorize-minorize algorithm. In particular, upon inspecting the proof of Proposition 4, the robust valuation maximizes over all  $V : \mathbb{D} \times \mathbb{H} \mapsto (0, +\infty)$ , the following function

$$f(V) = \mathbb{E}_{H \sim \mathbb{P}} \left[ R(H) \cdot \ln \left( \sum_{q=1}^{|H|} V(H_q \mid H_{[1, \dots, q-1]}) \right) - \sum_{q=1}^{|H|} V(H_q \mid H_{[1, \dots, q-1]}) \right].$$

An immediate consequence of Propositions 3 and 4 is that  $V$  maximizes  $f$  if and only if  $V$  is additive. Consequently,  $f$  can be thought of as a criterion to measure how well some valuation  $V$  satisfies the additivity property.



**Figure 3** Generative model for our synthetic experiments.

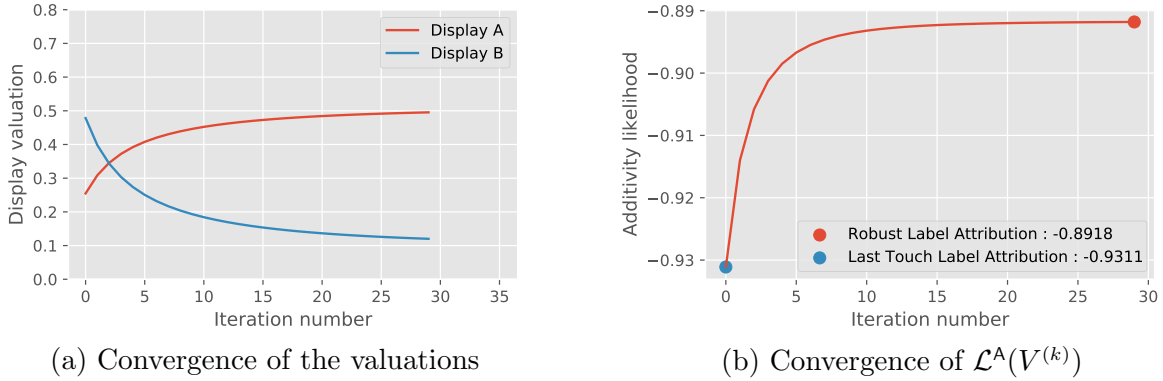
In practice, we do not know the reward function  $R$  but we are working with realized histories  $h^1, \dots, h^T$  and their associated rewards  $r^1, \dots, r^T$ . It is therefore natural to define the following likelihood for any valuation  $V$ ,

$$\mathcal{L}^A(V) = \frac{1}{T} \sum_t \left[ r^t \cdot \ln \left( \sum_{q=1}^{|h^t|} V(h_q^t | h_{[1, \dots, q-1]}^t) \right) - \sum_{q=1}^{|h^t|} V(h_q^t | h_{[1, \dots, q-1]}^t) \right]. \quad (5)$$

We use this likelihood to measure the progress of Algorithm 1. In particular, when we implement Step 2-a, Proposition 4 guarantees that our algorithm converges. On the other hand, we do not have any theoretical guarantees of convergence when implementing Step 2-b with a ML algorithm. In this case, we can use the likelihood defined in Equation (5) to check whether the algorithm converges and to evaluate the quality of the output. This is illustrated in Section 5.3.

## 5.2. Synthetic data

We generate a dataset inspired by the motivating example in Section 1. In particular, we assume that there are two types of displays. Displays of type A influence the user’s behavior and increase the conversion probability. On the other hand, displays of type B have very little influence on the conversion probability. However, they always appear right before a conversion. We capture the essence of this example by assuming that there are only two types of histories. The first type of history, denoted by  $h^1$ , consists of a single display A and happens with probability  $p = 1/3$ . The second type of history, denoted by  $h^2$ , consists of a display A followed by a display B and happens with probability  $1 - p$ . In the first case, a conversion happens with probability 0.5 and in the second case with probability 0.6. With our notation, this means that  $R(h^1) = 0.5$  and  $R(h^2) = 0.6$ . Figure 3 illustrates this generative procedure. Note that, unlike the theoretical framework presented in Section 3, we do not have access to the reward function  $R$  but only to histories and their realized rewards. In other words, our dataset contains histories, either  $h^1$  or  $h^2$ , and an associated reward of 1 or 0 depending on whether a conversion happens. We generate  $T = 10,000$  user histories  $h^1, \dots, h^T$  and denote by  $r^1, \dots, r^T$  their corresponding reward.

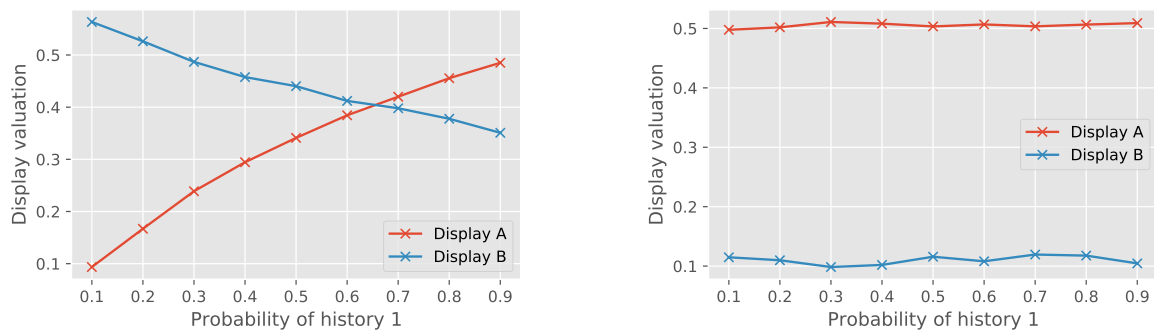


**Figure 4** Convergence of Algorithm 1 on synthetic data. The display valuation converges to the expected values of the generative model, and the additivity likelihood of the robust attribution  $\mathcal{L}^A(V^{(k)})$  monotonically converges and outperforms  $\mathcal{L}^A(V^{\text{LT}})$ .

We implement Algorithm 1 using this synthetic model<sup>2</sup>. We use the last touch label attribution to initialize the algorithm. For this synthetic model, we can either implement Step 2-a or Step 2-b of Algorithm 1. We begin by discussing the averaging implementation, i.e. Step 2-a, whose results are reported in Figure 4. Figure 4b shows that  $\mathcal{L}^A(V^{(k)})$  converges quickly, in about 20 iterations. Moreover, we see that the likelihood increases in each iteration and the final likelihood is higher than the initial point, i.e. the likelihood under the last touch valuation  $\mathcal{L}^A(V^{\text{LT}})$ . This is consistent with our theory. Furthermore, Figure 4a shows how the valuation  $V$  evolves for each type of display. Note that  $V(x_A|[]) \rightarrow 0.5$  and  $V(x_B|[x_A]) \rightarrow 0.1$  which is consistent with the underlying model parameters unlike  $V^{\text{LT}}$  (initial values in Figure 4a). This shows that our proposed method overcomes the weakness of the last touch method highlighted in the example from Section 1. We also illustrate the distributional robustness of our label attribution (as opposed to any other label attribution scheme) in Figure 5. In particular, we change the probability  $p$  used to generate the different history types while keeping the same reward function  $R$ . Figure 5 shows that our robust attribution is not affected by these changes in the generative process. On the other hand, the valuation computed using  $V^{\text{LT}}$  changes drastically: the ordering of the valuations between type A and type B displays even changes as  $p$  varies.

Interestingly, in this synthetic setting, using Step 2-b with a logistic regression implementation gives exactly the same results. Indeed, the setup is simple enough that the logistic regression is able to perfectly learn the model. Although, we do not report the results both implementations are available in the code.

<sup>2</sup>The code is available here: [https://colab.research.google.com/github/criteo-research/robust-label-attribution/blob/main/Robust\\_label\\_attribution\\_synthetic\\_dataset.ipynb](https://colab.research.google.com/github/criteo-research/robust-label-attribution/blob/main/Robust_label_attribution_synthetic_dataset.ipynb)



(a) Last touch attribution is not distributionally robust (b) Robust attribution is indeed distributionally robust

Figure 5 Illustration of distributional robustness guaranteed by Proposition 2.

### 5.3. “Criteo Attribution Modeling for Bidding” dataset

In this section, we apply our framework to a real dataset with a large feature space. The feature space size prohibits the use of Step 2-a in Algorithm 1, since each input is unique with very high probability, as it is often the case in ML problems. In this type of contexts, we often rely on ML techniques.

**5.3.1. Data description.** “The dataset represents a sample of 30 days of Criteo live traffic data. Each line corresponds to one impression (a banner) that was displayed to a user. For each banner we have detailed information about the context, if it was clicked, if it led to a conversion and if it led to a conversion that was attributed to Criteo or not. Data has been sub-sampled and anonymized so as not to disclose proprietary elements” (Diemert et al. 2017). In total, the dataset contains approximately 16M displays shown to 6M users among which 180K have converted. Each display has about 20 features including 10 categorical features describing contextual information. Some other notable features include whether the display was clicked or not and the time since the last click (see Diemert et al. (2017) for a complete description of the Criteo Attribution Modeling for Bidding Dataset).

**5.3.2. Implementation details.** We focus on displays that were clicked –which is a feature in the dataset– to have a stronger signal and be closer in spirit to the standard last touch benchmark (see Appendix E for more comprehensive approach). To mitigate the side effects caused by the fixed window of observation, we remove the histories for which some clicks are missing (such information can be inferred from the data).

In order to compare our results with existing metrics (see “Performance” paragraph in Section 5.3.3 below), we need binary rewards. Consequently, we split each history that triggered multiple conversions into multiple histories. The resulting dataset contains 16M displays shown to 8M users among which 196K have converted. The ML brick we use in Step 2-b of Algorithm 1 is

the combination of logistic regression and hashing trick described in Chapelle et al. (2014). The hashing trick (also referred to as feature hashing) is a machine learning strategy to embed the feature space into a vector of a controlled size, in our case  $2^{13}$ .

**5.3.3. Results and discussion.** For reproducibility purposes, the code of the experiment is available online<sup>3</sup>. We train our model using 80 % of the dataset and test on the remaining 20 %. The split is done on the user identifiers.

**Convergence.** We check the convergence of our algorithm using  $\mathcal{L}^A(V^{(k)})$  as a criterion. Figure 6a shows that the procedure converges in a few iterations similar to the previous set of experiments. Figure 6b also illustrates that our solution generalizes to the testing dataset. Note that in order to get this out of sample performance, we use a regularized logistic regression in Step 2-b of Algorithm 1. The downside is that it introduces a small bias as can be seen by the non-monotonicity of the iterates in Figure 6. Nevertheless, we observe a significant improvement compared to the last touch method.

To put these results in perspective, we additionally tested the sensitivity of our results with respect to the size of the feature space. Indeed, with the hashing trick, we are embedding our feature space into a target space whose size can be controlled. Figure 7 illustrates the value of the additivity likelihood  $\lim_k \mathcal{L}^A(V^{(k)})$  under different feature space sizes. As expected, the likelihood of the robust valuation increases with respect to the feature space size. On the other hand, there is a priori no reason to believe that when increasing the size of the feature space,  $\mathcal{L}^A(V^{\text{LT}})$  also increases. Indeed, the last touch method, unlike our robust valuation, is not designed to maximize  $\mathcal{L}^A$ . Somewhat reassuringly, we observe that  $\mathcal{L}^A(V^{\text{LT}})$  also increases with the feature space size. The striking result is that the improvements of our robust valuation over the last touch valuation are similar in magnitude to increasing the size of the feature space by a factor five.

**Qualitative insights.** Having shown that our method is improving the additivity likelihood, we next try to get some insights into how the proposed valuation is capturing some key elements of the problem. We discuss two important effects that have been discussed in the literature and recognised in practice.

The first one is the notion that the effect of a click decays over time. Consequently, it is less valuable to place an ad immediately after another one has been clicked. Several works have tackled this problem by making some assumptions on the underlying mechanism at play and hard coding this decaying effect (Chapelle et al. 2014, Diemert et al. 2017, Zhang et al. 2014). On the contrary, we do not hard code this relationship into our algorithm, yet our valuation is able to learn this

<sup>3</sup> [https://colab.research.google.com/github/criteo-research/robust-label-attribution/blob/main/Robust\\_label\\_attribution\\_Criteo\\_dataset.ipynb](https://colab.research.google.com/github/criteo-research/robust-label-attribution/blob/main/Robust_label_attribution_Criteo_dataset.ipynb)

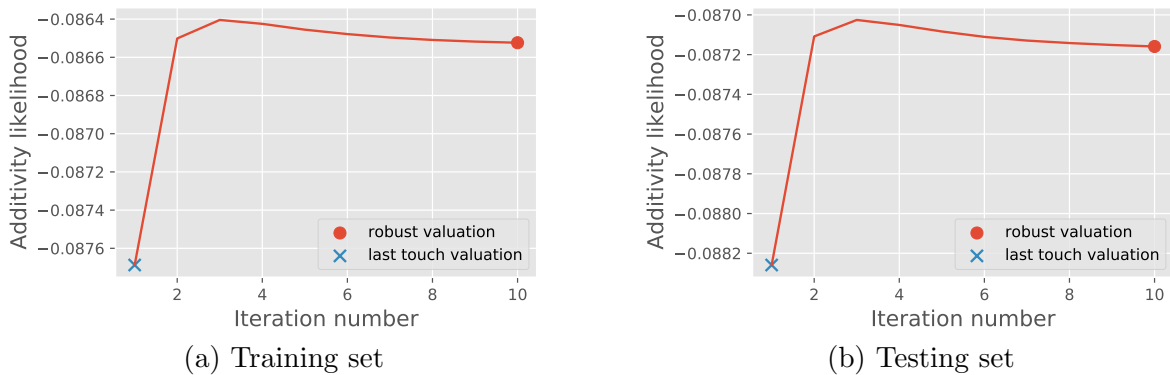


Figure 6 Convergence  $\mathcal{L}^A(V^{(k)})$  on the Criteo Attribution dataset.

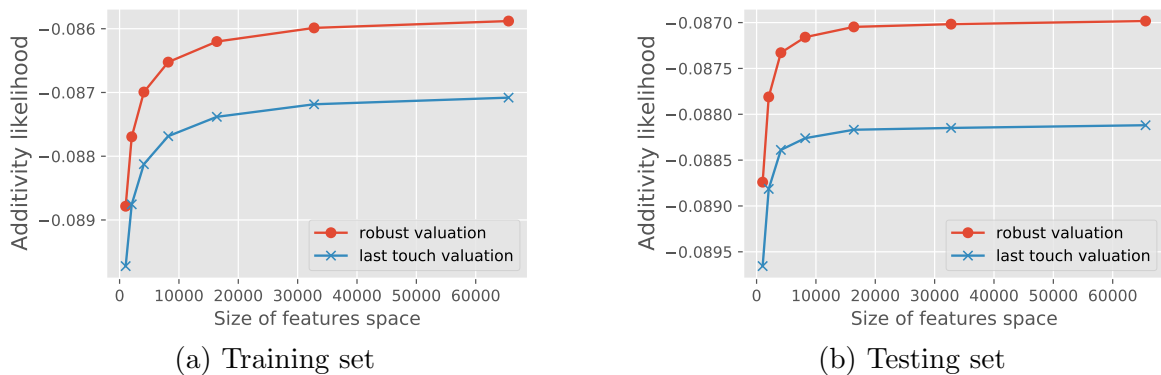
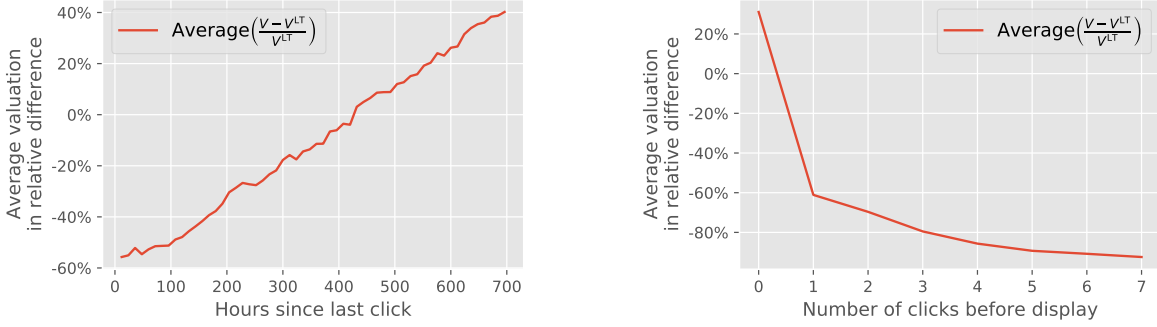


Figure 7 Additivity likelihood  $\mathcal{L}^A$  on the Criteo Attribution dataset as a function of the features space size.

effect from data. Indeed, we show in Figure 8a the relative difference between the robust and last touch valuation as a function of the feature “hours since last click”. We observe that the last touch valuation tends to overestimate the value of an opportunity when a display was just clicked and our robust valuation is able to correct for this effect. This is coherent with the live experiments done in Diemert et al. (2017).

Another effect that has been observed in practice is the diminishing marginal effect of ads. As a user is being exposed to different displays, the incremental benefit of showing an extra ad decreases. This is similar in spirit to the motivating example in the introduction. This is what we explore in Figure 8b where we show the relative difference between the robust and the last touch valuation as a function of the feature “number of clicks before display”. We observe that, compared to the last touch benchmark, the robust attribution framework drastically reduces the valuation of a display with the number of preceding clicks.

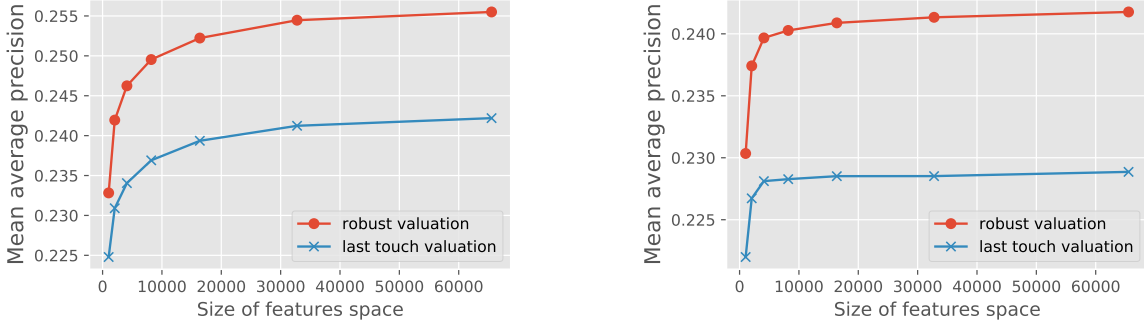
**Performance.** Finally, we also evaluate the performance of our approach with the procedure used in Zhang et al. (2014) for Multitouch Attribution models. This procedure consists in mapping the conversion probabilities of each display in a given history to a single history’s conversion probability



(a) Time elapsed since the last user's click

(b) Number of clicks already made by the user

**Figure 8** Relative variation of the display valuation as function of several features. The robust attribution tends to lower its valuation on recently exposed users.



(a) Training set

(b) Testing set

**Figure 9** Mean average precision for the history conversion probability given in Equation (6).

and was first introduced in Dalessandro et al. (2012) as a generative model. The proposed conversion probability of a history can be written with our notations as follows

$$\mathbb{P}(R^t = 1) = \left( 1 - \prod_{q=1}^{|h^t|} (1 - V(x_q^t | h_{[1, \dots, q-1]}^t)) \right) \cdot \delta^{|h^t|} \quad (6)$$

Here,  $\mathbb{P}(R^t = 1)$  denotes the probability that a history  $h^t$  leads to a conversion which is a binary reward, and for each display in the history  $h^t = [x_1^t, \dots, x_{|h^t|}^t]$ , we identify conversion probability of  $x_q^t$  with its valuation  $V(x_q^t | h_{[1, \dots, q-1]}^t)$ . The first term in brackets represents the probability that at least one display leads to a conversion assuming zero interaction effects. The second term  $\delta^{|h^t|}$  accounts for the marginally decreasing effect of each ad. Zhang et al. (2014) fixes  $\delta$  to 0.95. We use the same value for consistency. We can then compare two models by evaluating how well they rank the converted histories according to Equation (6). A natural metric in this case is the mean average precision metric since it summarises the Precision-Recall curve used in Zhang et al. (2014) with a single value. We present the results as a function of the size of the features space in Figure 9. While our model was designed to maximize the additivity likelihood, we observe that it also strongly out-performs the last touch valuation when using this different metric.

## 6. Conclusion

In this paper, we introduce the label attribution problem which consists of splitting the reward given to a bidder for each users into labels for each display opportunity. This is an important step needed to create a training dataset for most ML algorithm but it is often overlooked in practice. We formalize this problem and propose an approach to this problem which we call the robust label attribution. Our method is motivated by an intuitive additivity property and enjoys several theoretical structural properties. It is also practical and we show how it can be implemented at the scale required for display advertising. We end this paper by drawing some connections with related research topics and discuss some important future research directions.

**Connection to Shapley value.** The notion of Shapley value has been used for the advertiser attribution problem Singal et al. (2019). This is an appealing approach since the Shapley value of a coalition game is uniquely defined by four natural and desirable axioms. Note that our proposed solution satisfies three of these four axioms (linearity, null player and efficiency). It would be interesting to study if the concept of Shapley value could be applied in our setting and uncover connections with our proposed solution.

**Delay modeling.** One limitation of our approach is that it does not explicitly take into account the time delay between the display (or the click) and the sale, despite the fact that such delay might be informative on the relative contribution of each past events to the conversion: a 100 day-old display is less likely to cause a conversion than a more recent display. Several previous works have developed methodologies to capture time delay in a prediction framework (Zhang et al. 2014, Chapelle et al. 2014). It would be an interesting extension to combine these ideas with our robust label attribution methodology.

**Causality.** Our paper is related to causal statistics. Indeed, observe that if (a) we merge the advertiser and the bidders into a unique entity, (b) the reward is set as the sum of all conversions, (c) we add a null event at the beginning of each sequence, then the resulting bid corresponds to the incremental expected number of sales generated by the interaction.

## References

- Anderl E, Becker I, Von Wangenheim F, Schumann JH (2016) Mapping the customer journey: Lessons learned from graph-based online attribution modeling. *International Journal of Research in Marketing* 33(3):457–474.
- Arava SK, Dong C, Yan Z, Pani A, et al. (2018) Deep neural net with attention for multi-channel multi-touch attribution. arXiv preprint arXiv:1809.02230.
- Archak N, Mirrokni VS, Muthukrishnan S (2010) Mining advertiser-specific user behavior using adfactors. *Proceedings of the 19th international conference on World wide web*, 31–40.



- Arjovsky M, Bottou L, Gulrajani I, Lopez-Paz D (2019) Invariant risk minimization. arXiv preprint arXiv:1907.02893.
- Balseiro S, Kim A, Mahdian M, Mirrokni V (2021) Budget-management strategies in repeated auctions. *Operations Research* .
- Balseiro SR, Candogan O (2017) Optimal contracts for intermediaries in online advertising. *Operations Research* 65(4):878–896.
- Balseiro SR, Gur Y (2019) Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science* 65(9):3952–3968.
- Berman R (2018) Beyond the last touch: Attribution in online advertising. *Marketing Science* 37(5):771–792.
- Blake T, Nosko C, Tadelis S (2015) Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica* 83(1):155–174.
- Bompaire M, Gilotte A, Heymann B (2020) Causal models for real time bidding with repeated user interactions.
- Chapelle O (2014) Modeling delayed feedback in display advertising. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1097–1105.
- Chapelle O, Manavoglu E, Rosales R (2014) Simple and scalable response prediction for display advertising. *ACM Transactions on Intelligent Systems and Technology* 5(4).
- Choi H, Mela CF, Balseiro SR, Leary A (2020) Online display advertising markets: A literature review and future directions. *Information Systems Research* 31(2):556–575.
- Dalessandro B, Stitelman O, Perlich C, Provost F (2012) Causally motivated attribution for online advertising. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1–9.
- Danaher PJ (2018) Delusion in Attribution : Caveats in Using Attribution for Multimedia Budget Allocation. *Journal of Marketing Research* 55(April):667–685.
- Dhamdhere K, Sundararajan M, Yan Q (2018) How important is a neuron? arXiv preprint arXiv:1805.12233.
- Diemert E, Galland P, Meynet J, Lefortier D (2017) Attribution modeling increases efficiency of bidding in display advertising. *2017 AdKDD and TargetAd - In conjunction with the 23rd ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2017, 2* (ACM).
- Enberg J (2019) Global Digital Ad Spending 2019. URL <https://www.emarketer.com/content/global-digital-ad-spending-2019>.
- Flores-Szwagrzak K, Treibich R (2020) Teamwork and individual productivity. *Management Science* 66(6):2523–2544.
- Golrezaei N, Lobel I, Paes Leme R (2018) Auction design for roi-constrained buyers. *Available at SSRN 3124929* .

- Heymann B (2019) Cost per action constrained auctions. *Proceedings of the 14th Workshop on the Economics of Networks, Systems and Computation*, 1–8.
- Hunter DR, Lange K (2004) A tutorial on MM algorithms. *The American Statistician* 58(1):30–37.
- Ji W, Wang X (2017) Additional multi-touch attribution for online advertising. *31st AAAI Conference on Artificial Intelligence, AAAI 2017* 1360–1366.
- Ji W, Wang X, Zhang D (2016) A probabilistic multi-touch attribution model for online advertising. *International Conference on Information and Knowledge Management, Proceedings 24-28-October-2016*:1373–1382.
- Johnson GA, Lewis RA, Nubbemeyer EI (2017) Ghost Ads: Improving the economics of measuring online ad effectiveness. *Journal of Marketing Research* 54(6):867–884.
- Laub PJ, Taimre T, Pollett PK (2015) Hawkes Processes. Technical report.
- Lemon K (2016) Research Priorities 2016–2018. Technical report.
- Lewis RA, Mrao J, Reiley DH (2011) Here, there, and everywhere: Correlated online behaviors can lead to overestimates of the effects of advertising. *Proceedings of the 20th International Conference on World Wide Web, WWW 2011*, 157–166.
- Lewis RA, Rao JM (2015) The unfavorable economics of measuring the returns to advertising. *Quarterly Journal of Economics* 130(4):1941–1973.
- Perlich C, Dalessandro B, Hook R, Stitelman O, Raeder T, Provost F (2012) Bid optimizing and inventory scoring in targeted online advertising. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* .
- Peters J, Bühlmann P, Meinshausen N (2016) Causal inference by using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 78(5):947–1012.
- Shao X, Li L (2011) Data-driven multi-touch attribution models. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 258–264.
- Singal R, Besbes O, Desir A, Goyal V, Iyengar G (2019) Shapley meets uniform: An axiomatic framework for attribution in online advertising. *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019* 1713–1723.
- Wang J, Zhang W, Yuan S, et al. (2017) Display advertising with real-time bidding (RTB) and behavioural targeting. *Foundations and Trends in Information Retrieval* 11(4-5):297–435.
- Xu J, Shao X, Ma J, Lee Kc, Qi H, Lu Q (2016) Lift-based bidding in ad selection. *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.
- Xu L, Duan JA, Whinston A (2014) Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science* 60(6):1392–1412.

Zhang Y, Wei Y, Ren J (2014) Multi-touch Attribution in Online Advertising with Survival Theory. *Proceedings - IEEE International Conference on Data Mining, ICDM 2015-January*(January):687–696.

## Appendix

### Appendix A: Proof of Proposition 1

Fix a history  $h \in \mathbb{H}$  and display  $x \in \mathbb{D}$ . Recall that bidding  $V(x|h)$  is myopic weakly dominant if and only if in a second price setting, bidding  $V(x|h)$  is weakly dominant optimal for any competition profile. For a given competition profile, assume that the highest bid among competition follows a density  $g$ . The optimal bid  $V(x|h)$  must maximize the following

$$\int_0^{V(x|h)} (R(h + [x]) - t)g(t)dt + \int_{V(x|h)}^{+\infty} g(t)R(h)dt,$$

where the first integral corresponds to the set of competition prices where we win the auction and the second term correspond to competition bids above  $V$ . In the latter case, we obtain  $R(h)$  because of the missed opportunity. Writing the first order condition yields

$$(R(h + [x]) - V(x|h) - R(h))g(V(x|h)) = 0$$

for any  $g$ , hence, bidding  $V(x|h)$  is myopic weakly dominant if and only if  $R(h + [x]) - V(x|h) - R(h) = 0$ .

### Appendix B: Proof of Proposition 2

We assume for this proof that Proposition 3, whose proof is presented next, holds. By Proposition 3, if  $V$  is myopic optimal, then it is associated with the internal attribution given in Equation (2) which does not depend on  $\mathbb{P}$ . We therefore only have to prove the reverse implication which we do next.

Assume that  $V$  does not depend on  $\mathbb{P}$  but only on the reward function  $R$ , i.e. there exists an internal attribution  $\mu$  such that  $V = V_{\mathbb{P}}^{\mu} = V_{\mathbb{Q}}^{\mu}$  for any other probability distribution  $\mathbb{Q}$  over  $\mathbb{H}$ . We begin by showing that there exists a unique valuation which is distributionally robust, i.e., if we let  $\mu$  its associated internal attribution, we must have for any probability  $\mathbb{P}$  and  $h \in \mathbb{H}$

$$V(h) = V_{\mathbb{P}}^{\mu}(h) = \mathbb{E}_{H \sim \mathbb{P}}[\mu(|h|, H) | H \succcurlyeq h].$$

The internal attribution defined by Equation (2) is distributionally robust since the expression of  $\mu$  does not depend on  $\mathbb{P}$  which shows existence. We show unicity by contradiction. Assume that there are two such internal attributions  $\mu^1$  and  $\mu^2$ . For all  $\mathbb{P}$ ,  $h$  and  $j \leq |h|$ , we must have

$$V(h_j | h_{[1, \dots, j-1]}) = \mathbb{E}_{H \sim \mathbb{P}}[\mu^1(j, H) | H \succcurlyeq h_{[1, \dots, j]}] = \mathbb{E}_{H \sim \mathbb{P}}[\mu^2(j, H) | H \succcurlyeq h_{[1, \dots, j]}].$$

Consider a probability  $\mathbb{Q}$  such that the weight of all elements  $H \succcurlyeq h_{[1, \dots, j]}$  is moved to  $h_{[1, \dots, j]}$ . Then, we have for all  $h$  and  $j \leq |h|$ ,  $\mu^1(j, h) = \mu^2(j, h)$ . Note that since the probability  $\mathbb{Q}$  must satisfy Assumption 2, a limiting argument yields the result.

To conclude the proof, note that the unique associated valuation is myopic optimal by Proposition 3.

### Appendix C: Proof of Proposition 3

We first assume  $V$  satisfies additivity property from Definition 5. For all  $h \in \mathbb{H}$  such that  $R(h) > 0$  and  $i \in \{1, \dots, |h|\}$ , let

$$\mu(i, h) = \frac{V(h_i | h_{[1, \dots, i-1]})}{\sum_{j=1}^{|h|} V(h_j | h_{[1, \dots, j-1]})} \cdot R(h).$$

It is clear that  $\mu$  is a label attribution function as stated in Definition 3. Now, for all  $h \in \mathbb{H}$ , we have, using Equation (1),

$$\begin{aligned} V_{\mathbb{P}}^{\mu}(h_{|h|} | h_{[1, \dots, |h|-1]}) &= \mathbb{E}_{H \sim \mathbb{P}} \left[ \frac{V(H_{|h|} | H_{[1, \dots, |h|-1]})}{\sum_{j=1}^{|H|} V(H_j | H_{[1, \dots, j-1]})} R(H) \mid H \succcurlyeq h \right] \\ &= V(h_{|h|} | h_{[1, \dots, |h|-1]}) \cdot \mathbb{E}_{H \sim \mathbb{P}} \left[ \frac{R(H)}{\sum_{j=1}^{|H|} V(H_j | H_{[1, \dots, j-1]})} \mid H \succcurlyeq h \right] \\ &= V(h_{|h|} | h_{[1, \dots, |h|-1]}), \end{aligned}$$

where the last equality follows from the fact that  $V$  is additive.

We next show the reverse implication. We assume that the fixed point equation is verified

$$\mu(i, h) = \frac{V(h_i | h_{[1, \dots, i-1]})}{\sum_{j=1}^{|h|} V(h_j | h_{[1, \dots, j-1]})} \cdot R(h).$$

Using the associated valuation definition, we have for all  $(x, h) \in \mathbb{D} \times \mathbb{H}$ ,

$$V_{\mathbb{P}}^{\mu}(x|h) = \mathbb{E}_{H \sim \mathbb{P}} [\mu(|h| + 1, H) \mid H \succcurlyeq h + [x]].$$

Consequently, for all  $h \in \mathbb{H}$  and  $i \in \{1, \dots, |h|\}$ , we obtain

$$\begin{aligned} \sum_{j=1}^{|h|} V(h_j | h_{[1, \dots, j-1]}) &= \sum_{j=1}^{|h|} \mathbb{E}_{H \sim \mathbb{P}} [\mu(j, H) \mid H \succcurlyeq h_{[1, \dots, j]}] \\ &= \sum_{i=1}^{|h|} \mathbb{E}_{H \sim \mathbb{P}} \left[ \frac{V(h_i | h_{[1, \dots, i-1]})}{\sum_{j=1}^{|h|} V(h_j | h_{[1, \dots, j-1]})} \cdot R(h) \mid H \succcurlyeq h_{[1, \dots, i]} \right] \\ &= R(h), \end{aligned}$$

where we obtain de last equality by interchanging the expectation and sum operators. This concludes the proof.

### Appendix D: Proof of Proposition 4

To prove the convergence of the iterative procedure defined in Equations (3)-(4), we show that it corresponds to the iterates of a majorize-minorize (MM) algorithm. We remind the reader of the MM philosophy and refer to Hunter and Lange (2004) for more details. An MM algorithm aims at finding the maximizer  $\theta^*$  of a function  $f(\theta)$ . For that purpose, a surrogate function  $g(\theta|\hat{\theta})$  that depends on a current estimate  $\hat{\theta}$  and is typically easier to maximize is identified and maximized instead. The key is to choose a surrogate function that minorizes the objective function. A typical iteration of the MM algorithm can be written as follows. Given an estimate  $\theta^{(m)}$ , we let

$$\theta^{(m+1)} = \arg \max_{\theta} g(\theta|\theta^{(m)}).$$

The procedure is guaranteed to converge to a local maximum of  $f$  as long as the following are satisfied: (1)  $f(\theta)$  is concave, (2)  $g(\theta^{(m)}|\theta^{(m)}) = f(\theta^{(m)})$ , and (3)  $g(\theta|\theta^{(m)}) \leq f(\theta), \forall \theta$ . We show that our iterative procedure is a special case of an MM algorithm by letting for all  $V : \mathbb{D} \times \mathbb{H} \mapsto \mathbb{R}^+$ ,

$$f(V) = \mathbb{E}_{H \sim \mathbb{P}} \left[ R(H) \cdot \ln \left( \sum_{q \in \{1, \dots, |H|\}} V(H_q | H_{[1, \dots, q-1]}) \right) - \sum_{q \in \{1, \dots, |H|\}} V(H_q | H_{[1, \dots, q-1]}) \right],$$

$$g(V|\hat{V}) = \mathbb{E}_{H \sim \mathbb{P}} \left[ \sum_{j \in \{1, \dots, |H|\}} \left\{ \frac{\hat{V}(H_j | H_{[1, \dots, j-1]}) \cdot R(H)}{\sum_{q \in \{1, \dots, |H|\}} \hat{V}(H_q | H_{[1, \dots, q-1]})} \cdot \ln \left( \frac{V(H_j | H_{[1, \dots, j-1]})}{\hat{V}(H_j | H_{[1, \dots, j-1]})} \sum_{q \in \{1, \dots, |H|\}} \hat{V}(H_q | H_{[1, \dots, q-1]}) \right) \right. \right. \\ \left. \left. - V(H_j | H_{[1, \dots, j-1]}) \right\} \right].$$

Since the space of histories is finite,  $V$  can always be mapped to a finite dimension vector. We do however keep the function notation for simplicity. First, note that  $f$  is strictly concave as a linear combination of concave functions. Second, observe that for all  $V$ ,

$$g(V|V) = \mathbb{E}_{H \sim \mathbb{P}} \left[ \sum_{j \in \{1, \dots, |H|\}} \left\{ \frac{V(H_j | H_{[1, \dots, j-1]}) \cdot R(H)}{\sum_{q \in \{1, \dots, |H|\}} V(H_q | H_{[1, \dots, q-1]})} \cdot \ln \left( \frac{V(H_j | H_{[1, \dots, j-1]})}{V(H_j | H_{[1, \dots, j-1]})} \sum_{q \in \{1, \dots, |H|\}} V(H_q | H_{[1, \dots, q-1]}) \right) \right. \right. \\ \left. \left. - V(H_j | H_{[1, \dots, j-1]}) \right\} \right] \\ = \mathbb{E}_{H \sim \mathbb{P}} \left[ R(H) \cdot \ln \left( \sum_{q \in \{1, \dots, |H|\}} V(H_q | H_{[1, \dots, q-1]}) \right) \cdot \sum_{j \in \{1, \dots, |H|\}} \frac{V(H_j | H_{[1, \dots, j-1]})}{\sum_{q \in \{1, \dots, |H|\}} V(H_q | H_{[1, \dots, q-1]})} \right. \\ \left. - \sum_{j \in \{1, \dots, |H|\}} V(H_j | H_{[1, \dots, j-1]}) \right] \\ = f(V).$$

Third, the inequality  $f(V) \geq g(V|\hat{V})$  follows by using the concavity of the logarithm and Jensen inequality on the random variable  $X$  which take value

$$\frac{V(H_j | H_{[1, \dots, j-1]})}{\hat{V}(H_j | H_{[1, \dots, j-1]})} \cdot \sum_{q \in \{1, \dots, |H|\}} \hat{V}(H_q | H_{[1, \dots, q-1]})$$

with probability  $\hat{V}(H_j | H_{[1, \dots, j-1]}) / \sum_{q \in \{1, \dots, |H|\}} \hat{V}(H_q | H_{[1, \dots, q-1]})$ . Finally, for any  $\hat{V}$ , we have  $\arg \max_V g(V|\hat{V}) = V_{\mathbb{P}}^{\mu_{\hat{V}}}$  by Lemma 1 below. Using Hunter and Lange (2004), we conclude that the sequence of  $V^{(k)}$  corresponds to the iterates of an MM algorithm with strictly concave criteria, and hence converges. This concludes the proof of Proposition 4.

LEMMA 1. For any  $\hat{V} : \mathbb{H} \mapsto \mathbb{R}$ ,  $\arg \max_V g(V|\hat{V}) = V_{\mathbb{P}}^{\mu_{\hat{V}}}$ .

*Proof.* Since  $g(V|\hat{V})$  is concave in  $V$ , it suffices to show that  $V_{\mathbb{P}}^{\mu_{\hat{V}}}$  satisfies the first order condition. In particular, for any  $(x, h) \in \mathbb{D} \times \mathbb{H}$ ,

$$\frac{\partial g(V|\hat{V})}{\partial V(x|h)} = \frac{\partial}{\partial V(x|h)} \left( \sum_t \mathbb{P}(t) \cdot \sum_{j \in \{1, \dots, |t|\}} \left\{ \frac{\hat{V}(t_j | t_{[1, \dots, j-1]}) \cdot R(t)}{\sum_{q \in \{1, \dots, |t|\}} \hat{V}(t_q | t_{[1, \dots, q-1]})} \cdot \ln \left( \frac{V(t_j | t_{[1, \dots, j-1]})}{\hat{V}(t_j | t_{[1, \dots, j-1]})} \sum_{q \in \{1, \dots, |t|\}} \hat{V}(t_q | t_{[1, \dots, q-1]}) \right) \right. \right. \\ \left. \left. - V(t_j | t_{[1, \dots, j-1]}) \right\} \right)$$

Note that  $h$  appears in the inner sum if and only if  $t \succcurlyeq h + [x]$ . Therefore, we can write

$$\frac{\partial g(V|\hat{V})}{\partial V(x|h)} = \frac{\partial}{\partial V(x|h)} \left( \sum_{t \succcurlyeq h+[x]} \mathbb{P}(t) \cdot \left\{ \frac{\hat{V}(x|h) \cdot R(t)}{\sum_{q \in \{1, \dots, |t|\}} \hat{V}(t_q | t_{[1, \dots, q-1]})} \cdot \ln \left( \frac{V(x|h)}{\hat{V}(x|h)} \sum_{q \in \{1, \dots, |t|\}} \hat{V}(t_q | t_{[1, \dots, q-1]}) \right) - V(x|h) \right\} \right).$$

Exchanging the derivative and sum operators yields

$$\frac{\partial g(V|\hat{V})}{\partial V(x|h)} = \sum_{t \succcurlyeq h+[x]} \mathbb{P}(t) \cdot \left\{ \frac{\hat{V}(x|h) \cdot R(t)}{\sum_{q \in \{1, \dots, |t|\}} \hat{V}(t_q | t_{[1, \dots, q-1]})} \cdot \frac{1}{V(x|h)} - 1 \right\}.$$

Setting the above to zero for the first order condition then implies that for all  $(x, h) \in \mathbb{D} \times \mathbb{H}$ , we have

$$\begin{aligned} V(x|h) &= \frac{1}{\sum_{t \succcurlyeq h+[x]} \mathbb{P}(t)} \cdot \left( \sum_{t \succcurlyeq h+[x]} \mathbb{P}(t) \cdot \frac{\hat{V}(x|h) \cdot R(t)}{\sum_{q \in \{1, \dots, |t|\}} \hat{V}(t_q | t_{[1, \dots, q-1]})} \right) \\ &= \mathbb{E}_{H \sim \mathbb{P}} \left[ \frac{\hat{V}(x|h)}{\sum_{q \in \{1, \dots, |H|\}} \hat{V}(H_q | H_{[1, \dots, q-1]})} \cdot R(H) \mid H \succcurlyeq h + [x] \right] \\ &= V_{\mathbb{P}}^{\mu \hat{V}}(x|h). \end{aligned}$$

This concludes the proof.  $\square$

## Appendix E: Extension

In this section, we present a generalization of the framework presented in Section 3 that can capture more realistic settings. In particular, we introduce the notions of ex-ante myopic optimality and ex-post qualifiers. To motivate our extension, we begin with an example that highlights a key limitation of our robust label attribution algorithm.

### E.1. Buyers are clickers

For simplicity, assume that all conversions produce a reward from the advertiser. Consider a model where the bidder interacts with the user via identical displays but can also track whether the user clicks on the display or not. These events contain important information and should inform the label attribution. For instance, imagine, in an extreme case, that the population is made of two types of users. A user of type A, upon seeing the ad, clicks on the ad and converts with probability 1. On the other hand, a user of type B never clicks on any ads but may convert after seeing at least two ads with probability 0.1. We denote by  $\rho$  the proportion of users of type A in the population. Being able to capture this extra information (the fact that a display is clicked or not) is crucial in this example since it allows the bidder to identify the type of a user. We next explain why this information cannot be captured with the framework presented in Section 3.

Let  $x$  be the only possible display. The probability that the bidder gets a reward after just one display is the proportion of users of type A in the global population:  $R([x]) = \rho$ . The probability to get a reward after the second display is  $R([x, x]) = \rho + 0.1 \cdot (1 - \rho)$ . Using the robust label attribution without taking into account the click information, we obtain by the additivity property that we should bid  $V([x, x]) = 0.1 \cdot (1 - \rho)$  for the second display, no matter whether the first display was clicked or not. This is clearly not optimal

since the marginal contribution of the second display directly depends on whether or not the first display was clicked and ideally, the bid for the second display should be zero for users of type A.

The solution to this paradox is quite simple and intuitive: when evaluating the second display, one should take into account whether the first display was clicked or not. Note that the click event happens after the first display has been evaluated and consequently, our current framework cannot incorporate this extra information. We next present a more general framework that is able to capture this type of extra information. Note that we use a click event in the example but any other relevant event could also be captured.

## E.2. Ex-post qualifier and ex-ante valuation

In the following, we introduce two concepts that allow adapting the robust label attribution to a very large class of generative models. We assume that there exists a finite number of events that can occur between two displays. Examples of such event include “the display has been clicked” or “10 minutes elapsed since the display was shown”. Without loss of generality, since we can always redefine the set of events, we can assume that exactly one event occurs between two displays. We call those events *ex-post qualifiers* since they happen after the display they are associated with and denote the set of ex-post qualifiers by  $Q$ . We can now define a *qualified display* as a pair  $(x, q)$  where  $x \in \mathbb{D}$  and  $q \in Q$ . Let  $\mathbb{H}^Q$  be the set of nonempty sequences of elements of  $\mathbb{D} \times Q$ . A valuation  $V$  is now a function from  $\mathbb{D} \times Q \times \mathbb{H}^Q$  to  $\mathbb{R}^+$  such that for all  $(x, q, h) \in \mathbb{D} \times Q \times \mathbb{H}^Q$ ,  $V((x, q)|h)$  represents the expected value of the qualified display  $(x, q)$  given the history  $h$ .

If we apply the robust label attribution framework to these qualified displays, we recover the robust valuation  $V$  and the robust label attribution  $\mu$ . However, the valuation  $V$  cannot be used directly since we do not have access to the ex-post qualifiers at bidding time (we do not know if the display is going to be clicked or not). To account for possible ex-post qualifiers, we simply average on the possible realizations of the ex-post qualifier. We assume that we are given a probability distribution on the ex-post qualifiers. More precisely, for  $(x, q, h) \in \mathbb{D} \times Q \times \mathbb{H}^Q$ , let  $p(q|h, x)$  be the marginal probability that the display  $x$  is qualified by the ex-post qualifiers  $q$  conditioned on  $h$  when  $x$  is the last display.

**DEFINITION 6 (EX-ANTE VALUATION).** For any valuation  $V : \mathbb{D} \times Q \times \mathbb{H}^Q \rightarrow \mathbb{R}^+$ , let  $\tilde{V} : \mathbb{D} \times \mathbb{H}^Q \rightarrow \mathbb{R}^+$  be the ex-ante valuation associated to  $V$  defined for all  $(x, h) \in \mathbb{D} \times \mathbb{H}^Q$  as

$$\tilde{V}(x|h) \triangleq \mathbb{E}_{q \sim p} V((x, q)|h) = \sum_{q \in Q} p(q|h, x) \cdot V((x, q)|h) \quad (7)$$

## E.3. Generalized robust label attribution

We can adapt our main theoretical results as follow.

**COROLLARY 1.** *The following are equivalent:*

1. **Optimality:** *The ex-ante valuation  $\tilde{V}$  associated to  $V$  is myopic optimal*
2. **Additivity:** *For all  $h \in \mathbb{H}^Q$  and  $(x, q) \in \mathbb{D} \times Q$ ,*

$$V((x, q)|h) = R(h + [(x, q)]) - R(h), \quad (8)$$

3. **Distributional robustness:**  *$V$  does not depend on  $\mathbb{P}$  but only on the reward function  $R$ ,*
4. **Fixed-point characterization:**  *$V$  is the associated valuation of the label attribution  $\mu$  defined for all  $h$  and  $i \in \{1, \dots, |h|\}$  as*

$$\mu(i, h) = \frac{V(h_i|h_{[1, \dots, i-1]})}{\sum_{j=1}^{|h|} V(h_j|h_{[1, \dots, j-1]})} \cdot R(h), \quad (9)$$



**E.4. Revisiting the “Buyers are clickers” example.**

In this example, the robust valuation of a first display qualified by a click (resp. non clicked display) is 1 (resp. 0). Note that for the second display, there is no more qualifier so the robust valuation is equal to the ex-ante valuation. By the additivity property, if the first display was clicked, which we denote by the ex-post qualifier  $c$ , the robust valuation of the second display satisfies  $1 + V(x|(x, c)) = 1$ , and if the first display was not clicked  $0 + V(x|(x, \bar{c})) = 0.1$ . Hence, we obtain  $V(x|(x, c)) = 0$  and  $V(x|(x, \bar{c})) = 0.1$ .

Last, observe that denoting by  $p$  the probability that the user clicks on the first display, the ex-ante valuation of the first display is  $\tilde{V}(x) = 1 \cdot p + 0 \cdot (1 - p) = p$ .