



## Log-optimal anytime-valid E-values

Wouter M. Koolen<sup>a,\*</sup>, Peter Grünwald<sup>a,b</sup>

<sup>a</sup> Centrum Wiskunde & Informatica, Netherlands

<sup>b</sup> Leiden University, Netherlands



### ARTICLE INFO

#### Article history:

Received 20 February 2021

Received in revised form 15 September 2021

Accepted 16 September 2021

Available online 24 September 2021

#### Keywords:

E-value

Anytime-valid E-value

Hypothesis testing

Test martingales

Betting score

Implied target

### ABSTRACT

We consider the problem of measuring statistical evidence against a composite null hypothesis. We base our approach on the concept of an E-value, which measures evidence by the multiplication factor achieved by engaging in bets that are fair under the null. We adopt the log-optimality criterion for choosing among all possible E-values, which was considered earlier for a fixed sample size. We extend these ideas to sequential testing under optional stopping, by revisiting anytime-valid E-values. Our main contribution is the formulation of a sequential log-optimality criterion. We study its properties, and work out examples analytically and computationally.

© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

A core task in statistical testing is to measure evidence against a null hypothesis. The main contribution of this paper is to propose an optimality criterion for a certain class of evidence measures for sequential testing of composite nulls, study its properties theoretically and work out examples.

Our work takes as the starting point the concept of an E-value, which is a random variable measuring statistical evidence against a null hypothesis. E-values<sup>1</sup> are based on fair bets, and as such form the building blocks of statistical testing by means of Test Martingales, as studied by e.g. Shafer et al. [6], and also of the game-theoretic approach to probability and statistics advocated by Shafer and Vovk [7]. As summarised by Shafer [5], there is an ongoing discussion in the literature on the philosophy and methodology of statistical testing, the relative merits of E-values vs p-values for measuring evidence, and the practical advantage of E-values for communicating statistical results to laypeople in terms of the intuitive language of bets. Taking note of that interesting scientific discussion, here we take desirability of the E-value concept as given. Our main aim is to develop their theory for optimal *sequential* testing of *composite* null hypotheses.

Technically, E-valuehood is a constraint; it does not guide us which E-value we should pick. This selection question was studied for composite nulls by Grünwald et al. [2], who propose a criterion for obtaining powerful E-values. They call their construction GROW, for growth-rate optimal in the worst-case. The core idea is to optimise the expectation of the log-evidence under some alternative. Their construction is, by design, for a fixed sample size. It was not immediately clear how to extend their ideas to the sequential case. One piece of the puzzle was supplied by Ramdas et al. [3], who study the

\* Corresponding author.

E-mail addresses: [wmkoolen@cwi.nl](mailto:wmkoolen@cwi.nl) (W.M. Koolen), [pdg@cwi.nl](mailto:pdg@cwi.nl) (P. Grünwald).

<sup>1</sup> The name E-value is due to [10], see [2, Section 6] for historical context on the concept.

concept of anytime-valid E-values, where the sample size is not fixed but determined by some exogenous stopping rule.<sup>2</sup> The remaining missing step is to extend the optimality criterion to sequential testing. In this paper, we study how to obtain anytime-valid E-values that are similarly log-optimal.

This contribution is structured as follows. In Section 2 we review the definitions and motivate the new optimality criterion. In Section 3 we study the resulting optimal anytime-valid E-values by means of primal and dual characterisations. In Section 4 we work out concrete examples, and extract their lessons. We conclude with a modest discussion in Section 5.

## 2. Setup and problem motivation

In this section we will go over the required definitions, then formulate the new optimisation problem, and discuss a “game-theoretic” view.

### 2.1. Preliminaries

In this section we set up notation. We purposefully strip things down to the bare minimum, so that we can focus on conveying our ideas crisply. In particular, we delay all measure theoretic considerations to the discussion Section 5.

Let  $\mathcal{X}$  be an outcome space of interest. In the examples we will consider the binary case  $\mathcal{X} = \{0, 1\}$ , as well as the continuous case  $\mathcal{X} = \mathbb{R}$ . We will also fix a finite horizon  $T \geq 1$ . We will consider probability distributions on sequences  $\mathcal{X}^T$ , as represented either by a probability mass function or a density. We will often work with i.i.d. distributions, i.e.  $P(x^T) = \prod_{i=1}^T p(x_i)$  for some one-outcome distribution  $p$  on  $\mathcal{X}$ . We will also encounter Bayesian mixtures, i.e.  $P(x^T) = \int P_\theta(x^T)w(\theta) d\theta$  where  $w$  is a prior distribution over a suitably indexed family  $P_\theta$  of distributions.

We write  $\mathcal{X}^{\leq T} = \bigcup_{t=0}^T \mathcal{X}^t$  for the set of sequences of lengths up to  $T$ , and denote the zero-length sequence by  $\epsilon$ . A process  $E : \mathcal{X}^{\leq T} \rightarrow \mathbb{R}$  assigns a value to each sequence of each length. We represent a *randomised stopping time*<sup>3</sup> by a process  $\tau : \mathcal{X}^{\leq T} \rightarrow [0, 1]$ , where  $\tau(x^m)$  indicates the conditional probability of stopping directly after having seen prefix  $x^m$ . (We will impose  $\tau(x^T) = 1$  throughout to respect our time horizon  $T$ ). In particular, a distribution  $P$  on full-length sequences  $\mathcal{X}^T$  and a randomised stopping time  $\tau$  induce the stopped distribution  $P^\tau$  on prefixes  $\mathcal{X}^{\leq T}$ , which is given for each  $x^m \in \mathcal{X}^{\leq T}$  by

$$P^\tau(x^m) = P(x^m) \underbrace{\tau(x^m) \prod_{i=0}^{m-1} (1 - \tau(x^i))}_{\text{prob. of stop. at time} = m}, \tag{1}$$

where  $P(x^m) = \sum_{x_{m+1} \dots x_T} P(x^T)$  denotes the probability of observing the prefix  $x^m \in \mathcal{X}^{\leq T}$  according to the distribution  $P$  on  $\mathcal{X}^T$ . A *deterministic stopping time* is restricted to take values in  $\{0, 1\}$ . Our reason for studying randomised stopping times from the outset is twofold: they are closed under taking mixtures, which helps with arguments based on convex duality, and they are a more efficient representation for doing numerical computation (the alternative being explicitly maintaining a probability distribution over an enumeration of all deterministic stopping times).

With this background established, we now turn to our main object of study: the anytime-valid E-value.

**Definition 1.** We say that a non-negative process  $E \geq 0$  is an *anytime-valid E-value for distribution  $P$*  if  $\mathbb{E}_{x^m \sim P^\tau} [E(x^m)] \leq 1$  for any stopping time  $\tau$ . We further say that  $E$  is an *anytime-valid E-value for hypothesis class  $\mathcal{H}$*  if it is an anytime-valid E-value for every distribution  $P \in \mathcal{H}$ .

Note that, due to linearity of expectation, we may equivalently bound the expected value only for all *deterministic* stopping times. We will employ the following convenient notational shorthand. If  $E$  is a process and  $\tau$  is a randomised stopping time, then we denote by  $E_\tau$  the random variable that, under  $P$ , takes value  $E_\tau = E(x^m)$  with probability  $P^\tau(x^m)$ . In particular, this shortens the anytime-valid E-value requirement to  $\mathbb{E}_P [E_\tau] \leq 1$ .

*Type-1 error control.* Anytime-valid E-values measure evidence against the null in a way that is robust to external optional stopping. More precisely, fix an anytime-valid E-value  $E$  and error budget  $\alpha \in (0, 1)$ . Let’s assume that we receive a stopped sample  $x^m \sim P^\tau$ , drawn from an (unknown) combination of any  $P \in \mathcal{H}$  and randomised stopping time  $\tau$ . Suppose further that our decision is to reject the (entire) null  $\mathcal{H}$  when  $E(x^m) \geq \frac{1}{\alpha}$ . Then the probability that we falsely reject the null is at most  $\alpha$ . That is

<sup>2</sup> The terminology *anytime-valid E-values* used in this work emphasises the relation to the standard batch E-value. Ramdas et al. [3] adopt the name *safe E-process* to stress the contrast with non-negative (super)-martingales.

<sup>3</sup> Our randomised stopping times are called *behavioural stopping times* in the literature. A common and equivalent (see e.g. [8]) definition of randomised stopping times is as random deterministic stopping times, where a deterministic stopping time is a  $\{0, \dots, T\}$ -valued random variable such that the event  $\{\tau \leq t\}$  is known (measurable) at time  $t \leq T$ .

**Proposition 2** (Type-I error control). Let  $E$  be an anytime-valid E-value for null  $\mathcal{H}$ . For any confidence parameter  $\alpha \in (0, 1)$ , any distribution  $P \in \mathcal{H}$  and any randomised stopping time  $\tau$ , the false rejection probability is at most

$$\mathbb{P}_P \left\{ E_\tau \geq \frac{1}{\alpha} \right\} \leq \alpha \mathbb{E}_P [E_\tau] \leq \alpha.$$

**Proof.** The first inequality is Markov’s, while the second is the defining property of an anytime-valid E-value.  $\square$

We note that the error control holds for all stopping times, including “greedy” stopping at the first time  $t \in \{0, 1, \dots, T\}$  such that  $E(x^t) \geq \frac{1}{\alpha}$ . In particular, yet another way to present the same result is to claim that  $\frac{1}{E}$  is an anytime-valid p-value for  $\mathcal{H}$ , see e.g. [3].

Even though anytime-valid E-values are always safe, they may still be rather useless, like for instance the uninformative constant  $E(x^m) = 1$ . Next we turn to the question of obtaining useful anytime-valid E-values.

### 2.2. The log-optimal anytime-valid E-value

From now on, we fix an hypothesis class  $\mathcal{H}$  of interest, which we will call the *null*, and denote by  $\mathcal{E}_{\mathcal{H}}$  the set of anytime-valid E-values for  $\mathcal{H}$ . We will now be concerned with the selection of an anytime-valid E-value for disqualifying  $\mathcal{H}$ . To this end, we will choose an *alternative*  $(Q, \sigma)$  consisting of a distribution  $Q \notin \mathcal{H}$  and a randomised stopping time  $\sigma$ . We will interpret  $Q$  as encoding what we hope is a better explanation for the observations in case  $\mathcal{H}$  is false, and we will think of  $\sigma$  as a prediction of the stopping time our E-value may be subjected to in that case. Here we will assume that one<sup>4</sup> alternative  $(Q, \sigma)$  is chosen, and we will revisit making this choice in the discussion Section 5. This will allow us to define optimality.

**Problem 3** (Log-optimal anytime-valid E-value). Fix null  $\mathcal{H}$  and alternative  $(Q, \sigma)$ . The *log-optimal anytime-valid E-value* (LOAVEV) is the maximiser of the optimisation problem

$$\sup_{E \in \mathcal{E}_{\mathcal{H}}} \mathbb{E}_Q [\ln E_\sigma],$$

and, following [5], we call the maximum the *implied target*.

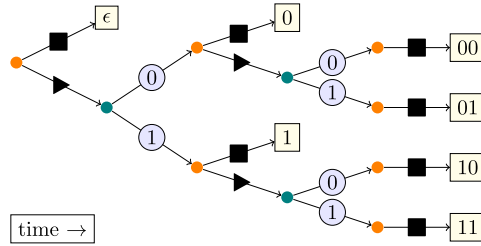
We now follow up with a brief motivation for this problem formulation, by discussing the problem that it aims to answer.

- We are looking for a measure of evidence against  $\mathcal{H}$  that is defined at every sample size.
- We want this measure to not report significant evidence against  $\mathcal{H}$  whenever the data are generated according to any  $P \in \mathcal{H}$ . This even has to hold under exogenous optional stopping. Not reporting significant evidence is formalised by the requirement that the expected value is at most one (see Proposition 2).
- Yet we want power under the alternative, i.e. when data come from  $Q$  and stopping is done according to  $\sigma$ . Our notion of power is the logarithmic growth rate. This has also been dubbed *implied target* by Shafer [5], who gives it an interpretation as a measure of the a-priori usefulness of performing a hypothesis test based on  $E$ .

The “power” objective  $\mathbb{E}_Q [\ln E_\sigma]$  takes its inspiration from the desire to do repeated testing, and can be motivated by a doubling rate argument (see e.g. [1, Section 6.1]). When  $n$  replications of the entire testing setup are drawn i.i.d. from  $Q^\sigma$ , then  $\frac{1}{n} \sum_{i=1}^n \ln E_\sigma^i \rightarrow \mathbb{E}_Q [\ln E_\sigma]$  by the weak law of the large numbers. Hence the evidence against the null accumulated by  $E$  over  $n$  interactions with  $Q^\sigma$  is asymptotically equivalent to  $e^{n \mathbb{E}_Q [\ln E_\sigma]}$  almost surely. The LOAVEV is defined to maximise this rate of accumulation. (See also the discussion in Section 5, where we investigate the power and (subtle, implicit) assumptions of this rate-based argument.)

After putting forward this problem, we need to characterise and analyse its solutions. In that respect, the contribution of this paper is twofold. First, we carefully work out the Lagrange dual problem to Problem 3. This dual form gives insight into the makeup of LOAVEVs, and it also presents opportunities for efficient numerical implementation. We then perform a sequence of analytical and numerical computations, answering questions about the makeup of optimal solutions. The main thrust is that we are able to solve the problem analytically in (a few) simple instances, and we can solve it numerically in more (small, finite) instances. Beyond these examples, the problem remains ill understood, and presents great opportunities for future exploration. We conclude this section with an illustration of the explicit temporal interaction in the LOAVEV problem.

<sup>4</sup> If we are entertaining several models  $Q^\sigma$ , we may merge them into a single one by taking a Bayesian mixture.



**Fig. 1.** Tree of situations for  $T = 2$  time steps and binary outcomes. Outcomes are displayed in blue circles, while the final sequence output is displayed in yellow boxes. The transition probabilities from a chance node (teal) are given by the conditional probabilities from the data-generating distribution  $P$  (or  $Q$ ). The transition probabilities from an optional stopping node (orange) are given by the randomised stopping time  $\tau$  (or  $\sigma$ ). The LOAVEV  $E$  from Problem 3 is a specific assignment of values to the yellow terminal nodes. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

### 2.3. A game-theoretic picture

We include a visualisation in Fig. 1 to make Problem 3 even more intuitive, concrete and explicit, and to illustrate the point that our Problem 3 is already interesting in the most basic setup: a binary alphabet  $\mathcal{X}$  and sample size  $T = 2$ . In the figure, the orange nodes represent optional stopping decisions, where one must either stop, resulting in a yellow terminal state labelled by the sequence of outcomes observed thus far, or one must continue. Upon continuing, one enters a teal node. The branching from each teal node models observing the next outcome, which is labelled in a blue circle.

Problem 3 asks to optimise over any-time E-values  $E$ , which we can represent as labelling each yellow terminal node (or equivalently, the orange node right before it) with a non-negative real number. The constraint  $E \in \mathcal{E}_{\mathcal{H}}$  asks that when an opponent controls the stopping decisions at the orange nodes, and when any element  $P \in \mathcal{H}$  of the null controls all the conditional distributions  $P(x_{m+1} | x^m)$  of the next outcome in each teal node, then the expected value of the terminal state reached is at most one. In addition, when the stopping decisions at the orange nodes are made by our stopping time  $\sigma$ , and the conditional distributions of outcomes are those of the alternative  $Q$ , we ask to maximise the expected logarithm of the value of the yellow terminal state reached.

### 3. Equivalence results

We start with a representation theorem. A process  $M$  is called a *test martingale* for  $P$  if it is a non-negative martingale (meaning that  $\mathbb{E}_P[M(x^{m+1}) | x^m] = M(x^m)$  for all  $x^m \in \mathcal{X}^{<T}$ ) starting from  $M(\epsilon) = 1$ . Let  $\mathcal{M}^P$  be the set of all test martingales for  $P$ . Then we have

**Lemma 4** (Primal representation).

$$\sup_{E \in \mathcal{E}_{\mathcal{H}}} \mathbb{E}_Q[\ln E_{\sigma}] = \sup_{(M^P)_{P \in \mathcal{H}} \in (\mathcal{M}^P)_{P \in \mathcal{H}}} \mathbb{E}_{x^m \sim Q^{\sigma}} \left[ \ln \inf_{P \in \mathcal{H}} M^P(x^m) \right].$$

The supremum ranges over collections of test martingales, one for each  $P \in \mathcal{H}$ .

**Proof.** If  $E$  is an anytime-valid E-value for  $\mathcal{H}$ , then  $E$  is dominated by a test martingale, say  $E \leq M^P$ , for each  $P$ , as shown by Ramdas et al. [3, Corollary 24, item 2]. Conversely,  $\inf_{P \in \mathcal{H}} M^P$  is an anytime-valid E-value because for every  $P \in \mathcal{H}$  we have  $E_P[\inf_{P' \in \mathcal{H}} M_{\tau}^{P'}] \leq E_P[M_{\tau}^P] = 1$ .  $\square$

What is the point of this Lemma? Well, the constraint of being in  $\mathcal{E}_{\mathcal{H}}$  requires checking something for *each* stopping time. This is both mathematically and computationally unwieldy. On the other hand, being a  $P$ -martingale is a simple check for each context  $\mathcal{X}^{<T}$ . In the finite-outcome case this reduces the number of variables from doubly exponential in  $T$  to merely singly exponential.

Next we turn to our main theorem. Let  $\text{KL}$  denote the Kullback-Leibler divergence, i.e.  $\text{KL}(Q \| P) = \sum_{x^m \in \mathcal{X}^{\leq T}} Q(x^m) \times \ln \frac{Q(x^m)}{P(x^m)}$ . Let  $\mathcal{B}$  be the set of distributions on  $\mathcal{X}^{\leq T}$  that can be represented as Bayesian mixtures of stopped elements of  $\mathcal{H}$ , i.e.  $\bar{P} \in \mathcal{B}$  iff  $\bar{P} = \mathbb{E}_{(P, \tau) \sim w} [P^{\tau}]$  for some joint prior distribution  $w$  on  $P \in \mathcal{H}$  and randomised stopping times  $\tau$ . We then have (proof in Appendix A)

**Theorem 5** (Primal-dual equivalence). For finite  $\mathcal{X}$ ,

$$\sup_{E \in \mathcal{E}_{\mathcal{H}}} \mathbb{E}_Q[\ln E_{\sigma}] = \inf_{\bar{P} \in \mathcal{B}} \text{KL}(Q^{\sigma} \| \bar{P}).$$

The restriction to a finite alphabet can be relaxed, though then some regularity conditions on  $\mathcal{H}$  do need to be imposed. From the proof it will be clear that a certain minimax result suffices (which indeed holds in case  $\mathcal{X}^{\leq T}$  is finite).

The upshot of this theorem is that instead of searching over anytime-valid E-values (or, equivalently, collections of test-martingales), we may perform a “Reverse Information Projection” (where “reverse” refers to minimising the KL in its second argument). Namely, we search among Bayesian mixtures (where the mixture is jointly over distributions  $P \in \mathcal{H}$  and stopping times  $\tau$ ) for the closest marginal to the alternative  $Q^\sigma$ .

*Reduction to batch viewpoint.* We may regard Theorem 5 as a specific reduction to the batch case considered by [2]. Here is how that works. The alternative  $(Q, \sigma)$  encodes a distribution  $Q^\sigma$  on  $\mathcal{X}^{\leq T}$ . We may also similarly reinterpret the null as the set of distributions on  $\mathcal{X}^{\leq T}$  given by

$$\mathcal{H}' := \{P^\tau \mid P \in \mathcal{H} \text{ and } \tau \text{ a randomised stopping time}\}.$$

In this viewpoint of treating sampling an outcome from  $X^{\leq T}$  as a single experiment, our main Problem 3 can be seen to reduce to the batch problem on outcome space  $\mathcal{X}^{\leq T}$  with alternative  $Q^\sigma$  and null  $\mathcal{H}'$ . It should be noted that the importance of this reduction is currently limited, as no general solution to the batch problem, either in analytic or numerical sense is as of yet available. Moreover, the blowup in size of  $\mathcal{H}'$  compared to  $\mathcal{H}$  may be substantial (there are doubly exponentially many stopping times, i.e. about  $2^{|\mathcal{X}^{T-1}|}$ ), making a computational approach based on this reduction prohibitive.

To aid in computation (and understanding), we propose the following alternative parametrisation. (We will not prove this proposition, instead we prove a stronger claim in the proof of Lemma 7 below.)

**Proposition 6** (Flow representation of stopped distributions). *Fix a distribution  $P$  on  $\mathcal{X}^T$ . The following two claims define and establish a bijection:*

- Let  $\tau$  be a randomised stopping time. Let  $\phi : \mathcal{X}^{<T} \rightarrow [0, 1]$  be defined by

$$\phi(x^m) := P\{X^m = x^m \text{ and } \tau > m\} = P(x^m) \prod_{i=0}^m (1 - \tau(x^i)).$$

That is,  $\phi(x^m)$  is the probability that a draw  $X^\tau \sim P^\tau$  has proper prefix  $x^m$ . Then

$$1 \geq \phi(\epsilon), \tag{2a}$$

$$\forall 1 \leq m \leq T, \forall x^m \in \mathcal{X}^m : \phi(x^{m-1})P(x_m \mid x^{m-1}) \geq \phi(x^m)\mathbf{1}_{m < T}. \tag{2b}$$

- Conversely, let  $\phi : \mathcal{X}^{<T} \rightarrow [0, 1]$  satisfy (2). Then

$$P_\phi(x^m) = \begin{cases} 1 - \phi(\epsilon) & m = 0 \\ \phi(x^{m-1})P(x_m \mid x^{m-1}) - \phi(x^m)\mathbf{1}_{m < T} & m > 0 \end{cases}$$

equals  $P^\tau$  for some randomised stopping time  $\tau$ .

One way to think about  $\phi$  is indicating the probability of visiting each teal node in Fig. 1 (note that we may indeed identify teal nodes with  $\mathcal{X}^{<T}$ ).

The usefulness of this parametrisation is that it allows mixing over  $P$ , while keeping the problem concave. We obtain the following more concise problem (with proof in Appendix B):

**Lemma 7** (Dual representation). *Let  $Q$  and  $P \in \mathcal{H}$  be distributions on sequences  $\mathcal{X}^T$ . Then the value*

$$\inf_{\bar{P} \in \mathcal{B}} \text{KL}(Q^\sigma \parallel \bar{P})$$

is equal to the minimum w.r.t.  $\phi : \mathcal{H} \times \mathcal{X}^{<T} \rightarrow [0, 1]$  of the function

$$\sum_{1 \leq m \leq T, x^m} Q^\sigma(x^m) \ln \frac{Q^\sigma(x^m)}{\sum_{P \in \mathcal{H}} (\phi(P, x^{m-1})P(x_m \mid x^{m-1}) - \phi(P, x^m)\mathbf{1}_{m < T})}$$

subject to the constraints

$$\sum_{P \in \mathcal{H}} \phi(P, \epsilon) = 1 - Q^\sigma(\epsilon),$$

$$\forall P \in \mathcal{H}, 1 \leq m \leq T, x^m : \phi(P, x^{m-1})P(x_m \mid x^{m-1}) \geq \phi(P, x^m)\mathbf{1}_{m < T}.$$

The two-argument function  $\phi$  in the above lemma encodes a joint distribution over  $P \in \mathcal{H}$  and randomised stopping times  $\tau$ . The way to think about this is that  $\phi(P, x^m)$  is the probability of using hypothesis  $P \in \mathcal{H}$ , and generating a sequence having  $x^m$  as a proper prefix (see Proposition 6). With this interpretation, indeed  $\phi(P, x^{m-1})$  is the probability of seeing  $x^{m-1}$  from  $P$  and not stopping yet, so that  $\phi(P, x^{m-1})P(x_m|x^{m-1})$  is the probability of seeing at least  $x^m$  from  $P$ . Then as  $\phi(P, x^m)$  is the probability of seeing strictly more than  $x^m$  from  $P$ , we have that  $\phi(P, x^{m-1})P(x_m|x^{m-1}) - \phi(P, x^m)$  is the probability of seeing exactly  $x^m$  from  $P$ . Finally, by summing over  $P$ , we find that  $\sum_{P \in \mathcal{H}} (\phi(P, x^{m-1})P(x_m|x^{m-1}) - \phi(P, x^m))$  is the probability of seeing exactly  $x^m$ .

Taking stock, in Lemma 7 we obtained a minimisation problem that is concave in the function  $\phi$ , which is itself specified by  $|\mathcal{H}| \cdot \frac{|\mathcal{X}|^t - 1}{|\mathcal{X}| - 1}$  many variables. This is “only” singly exponential in the sample size. While still exponential, this reduction allows us to computationally explore non-trivial examples in Section 4.

### 4. Examples

We consider five examples. First we look at the simple cases of singleton nulls, followed by deterministic stopping times. Then we look at a null of i.i.d. zero-mean Gaussians, where we are able pin down the LOAVEV analytically, and we study a generalisation based on properties of a reverse information projection. The results are pleasing and simple in these four cases, yet they do not illustrate the general complexity of the LOAVEV problem, which is, after all, the main focus of this paper. To study the latter, we work out a further numerical example at two sample sizes.

#### 4.1. Singleton $\mathcal{H} = \{P\}$

Suppose we want to be safe w.r.t a singleton  $\mathcal{H} = \{P\}$ . Then what is the LOAVEV? We claim it is the likelihood ratio  $E_t = \frac{Q_t}{P_t}$  regardless of  $\sigma$ . But why? And what does this reveal about the  $|\mathcal{H}| > 1$  case?

**Theorem 8.** Fix alternative  $(Q, \sigma)$  and point null  $\mathcal{H} = \{P\}$ . The likelihood ratio  $E_t = \frac{Q_t}{P_t}$  is LOAVEV (regardless of the stopping time  $\sigma$ ).

**Proof.** Let  $E_t = \frac{Q_t}{P_t}$  be the likelihood ratio test martingale, and let  $Z_t$  be any other  $P$ -test-martingale. Using the primal representation Lemma 4, to establish optimality of  $E_t$ , we need to show that

$$\mathbb{E}_{Q^\sigma} [\ln E_\sigma] \geq \mathbb{E}_{Q^\sigma} [\ln Z_\sigma].$$

Toward this goal, let  $f(\alpha) = \mathbb{E}_{Q^\sigma} [\ln((1 - \alpha)E_\sigma + \alpha Z_\sigma)]$ . We will show that  $f$  is maximised at  $\alpha = 0$ . Since  $f$  is concave, it suffices to check that  $f'(0) \leq 0$ . We have

$$f'(0) = \mathbb{E}_{Q^\sigma} \left[ \frac{Z_\sigma - E_\sigma}{E_\sigma} \right] = \mathbb{E}_{Q^\sigma} \left[ \frac{Z_\sigma}{\frac{Q_\sigma}{P_\sigma}} \right] - 1 \leq \mathbb{E}_{P^\sigma} [Z_\sigma] - 1 \leq 0,$$

where the middle inequality arises due to  $P^\sigma$  possibly putting mass where  $Q^\sigma$  does not, and the last inequality uses that  $Z_t$  is a test martingale for  $P$ . This proves that the likelihood ratio  $E_t$  is LOAVEV.  $\square$

#### 4.2. Deterministic stopping time $\sigma$

In the previous section, we characterised the LOAVEV for the case of point nulls, and we saw that the randomised stopping time  $\sigma$  of the alternative is immaterial. In this section we look at another special case, namely that of alternatives with a deterministic stopping time  $\sigma$ . We show that the anytime-validity requirement evaporates from the problem, and we are back in the batch setting of Grünwald et al. [2] with the sample size  $n$  replaced by  $\sigma$ .

**Corollary 9.** For alternative  $(Q, \sigma)$  with deterministic stopping time  $\sigma$ , we have

$$\sup_{E \in \mathcal{E}_\mathcal{H}} \mathbb{E}_Q [\ln E_\sigma] = \inf_{w \text{ prior on } \mathcal{H}} \text{KL} \left( Q^\sigma \parallel \int P^\sigma w(dP) \right).$$

**Proof.** Starting from Theorem 5, we see that the worst-case distribution is supported only on  $\tau = \sigma$ , since moving any mass from  $\tau \neq \sigma$  to  $\sigma$  strictly improves the objective.  $\square$

4.3. Null is normal with zero mean and unknown variance

In this section we denote by  $\mathcal{N}(\mu, \xi^2)$  the Gaussian distribution with mean  $\mu$  and variance  $\xi^2$ . As our null, we take the set of all zero-mean Gaussians  $\mathcal{H} = \{\mathcal{N}(0, \xi^2) \mid \xi^2 > 0\}$  and as our point alternative  $Q = \mathcal{N}(\mu, \rho^2)$  we take the fixed mean  $\mu \neq 0$  and variance  $\rho^2$ . We first characterise the LOAVEV.

**Theorem 10.** *The LOAVEV is the likelihood ratio process*

$$E(x^n) = \frac{P_{\mu, \rho^2}(x^n)}{P_{0, \rho^2 + \mu^2}(x^n)} = \left(\frac{\mu^2 + \rho^2}{\rho^2}\right)^{\frac{n}{2}} \exp\left(\frac{U_n}{2(\mu^2 + \rho^2)} - \frac{U_n - 2\mu S_n + n\mu^2}{2\rho^2}\right), \tag{3}$$

where  $S_n = \sum_{i=1}^n x_i$  and  $U_n = \sum_{i=1}^n x_i^2$ .

**Proof.** We will build on Theorem 5, by showing that the minimiser of the projection of  $\mathcal{N}(\mu, \rho^2)$  onto the null  $\mathcal{H}$  is the prior  $\delta_{\rho^2 + \mu^2}$  that puts all mass in the point  $\mathcal{N}(0, \rho^2 + \mu^2)$ . To this end, fix any other prior  $w$  on  $\xi^2$ . Let  $w_\alpha = \alpha w + (1 - \alpha)\delta_{\rho^2 + \mu^2}$ , and let  $U = \sum_{i=1}^n x_i^2$ . We will show that

$$-\mathbb{E}_{x^n \sim \mathcal{N}(\mu, \rho^2)} \left[ \ln \int w_\alpha(\xi^2) (2\pi \xi^2)^{-n/2} e^{-\frac{U}{2\xi^2}} d\xi^2 \right] \tag{4}$$

is maximised at  $\alpha = 0$ . To do this, let's check the derivative at  $\alpha = 0$ . We have

$$\frac{\partial}{\partial \alpha} (4) \Big|_{\alpha=0} = 1 - \mathbb{E}_{x^n \sim \mathcal{N}(\mu, \rho^2)} \left[ \frac{\int w(\xi^2) (2\pi \xi^2)^{-n/2} e^{-\frac{U}{2\xi^2}} d\xi^2}{\int w_0(\xi^2) (2\pi \xi^2)^{-n/2} e^{-\frac{U}{2\xi^2}} d\xi^2} \right],$$

which equals

$$1 - \mathbb{E}_{x^n \sim \mathcal{N}(\mu, \rho^2)} \left[ \int w(\xi^2) \frac{(\xi^2)^{-n/2}}{(\rho^2 + \mu^2)^{-n/2}} e^{\frac{U}{2(\rho^2 + \mu^2)} - \frac{U}{2\xi^2}} d\xi^2 \right].$$

Swapping the expectations, and resolving the expectation over  $U$  gives

$$1 - \int w(\xi^2) \frac{(\xi^2)^{-n/2}}{(\rho^2 + \mu^2)^{-n/2}} \frac{e^{\frac{a\mu^2}{2(1-a\rho^2)}}}{(1-a\rho^2)^{n/2}} d\xi^2 \quad \text{where } a = \frac{1}{\mu^2 + \rho^2} - \frac{1}{\xi^2}.$$

The integrand above is maximised at  $\xi^2 = \rho^2 + \mu^2$ , where it takes value 1. This proves the displayed expression is non-negative, establishing the desired optimality.  $\square$

It hence turns out that a likelihood ratio between the alternative  $\mathcal{N}(\mu, \rho^2)$  and the closest element of the null,  $\mathcal{N}(0, \rho^2 + \mu^2)$  is the log-optimal anytime-valid E-value (LOAVEV). We next show something stronger, namely that it is a test super-martingale for every element of the null.

**Lemma 11.** *The LOAVEV  $E$  from (3) is a test super-martingale (non-negative super-martingale starting from 1) under every distribution  $\mathcal{N}(0, \xi^2) \in \mathcal{H}$  from the null.*

**Proof.** Non-negativity and initial value 1 are clear. We have to show that the expected multiplicative increment is  $\leq 1$  for any variance  $\xi^2$ . We have

$$\frac{E(x^{n+1})}{E(x^n)} = \sqrt{\frac{\mu^2 + \rho^2}{\rho^2}} \exp\left(\frac{x_{n+1}^2}{2(\mu^2 + \rho^2)} - \frac{(x_{n+1} - \mu)^2}{2\rho^2}\right),$$

and hence the expected multiplicative increment is

$$\begin{aligned} & \frac{\mathbb{E}_{x_{n+1} \sim \mathcal{N}(0, \xi^2)} [E(x^{n+1}) \mid x^n]}{E(x^n)} \\ &= \sqrt{\frac{\mu^2 + \rho^2}{\rho^2}} \mathbb{E}_{x_{n+1} \sim \mathcal{N}(0, \xi^2)} \left[ e^{\frac{1}{2} \left( \frac{x_{n+1}^2}{\mu^2 + \rho^2} - \frac{(x_{n+1} - \mu)^2}{\rho^2} \right)} \mid x^n \right] \\ &= \frac{(\mu^2 + \rho^2)}{\sqrt{\mu^2(\rho^2 + \xi^2) + \rho^4}} \exp\left(-\frac{\mu^2(\mu^2 + \rho^2 - \xi^2)}{2(\mu^2(\rho^2 + \xi^2) + \rho^4)}\right). \end{aligned}$$

This expression is quasi-concave in  $\xi^2$ , and it is maximised by cancelling the derivative, revealing that the maximiser is  $\xi^2 = \rho^2 + \mu^2$  where the value is 1. This proves that  $E(x^n)$  is a test super-martingale for (every element of) the entire null  $\mathcal{H}$ .  $\square$

We conclude this example by contrasting the result with another, intuitive anytime-valid E-value. Namely, the likelihood ratio of the alternative and the maximum likelihood element of the null (which is at the empirical second moment  $\hat{\xi}^2 = \frac{U_n}{n}$ ). That is, we are looking at

$$E^{\text{ml}}(x^n) := \frac{P_{\mu, \rho^2}(x^n)}{\max_{\xi^2} P_{0, \xi^2}(x^n)} = \left(\frac{U_n}{n\rho^2}\right)^{n/2} e^{-\frac{U_n - 2\mu S_n + n(\mu^2 - \rho^2)}{2\rho^2}}.$$

This is an anytime-valid E-value. This can be seen, for example, by the fact that it is below the likelihood ratio with  $\xi^2$ , which is a test martingale for  $\xi^2$ . By construction,  $E^{\text{ml}}$  has a lower implied target than the LOAVEV. Let’s compare for a fixed  $n$ . Then the LOAVEV implied target (value of the LOAVEV problem) is the pleasantly simple

$$\frac{n}{2} \ln \frac{\mu^2 + \rho^2}{\rho^2},$$

while the implied target  $\mathbb{E}_Q[\ln E_n^{\text{ml}}]$  of the ML E-value unfortunately does not admit a closed form expression. We can evaluate it numerically for specific inputs, for example at  $\mu = 1$ ,  $\rho^2 = 1$  and  $n = 10$  we find implied targets 3.46574 and 3.07214. So indeed the ML-based E-value is slightly worse.

#### 4.4. Cases where the reverse information projection is a point

In the Gaussian case we found that the LOAVEV is a likelihood-ratio-based test super-martingale. We now identify a more general condition under which the same happens. Throughout this section, we will be working with i.i.d. distributions in the null and alternative. We will keep denoting  $T$ -outcome i.i.d. distributions by capital letters ( $P/Q/\dots$ ) and denote the corresponding one-outcome distributions by the corresponding small letters ( $p/q/\dots$ ). We will further denote by  $\mathcal{P}$  the set of one-outcome distribution whose i.i.d. extensions comprise the null  $\mathcal{H} = \{P \mid P \text{ is i.i.d. } p \text{ for } p \in \mathcal{P}\}$ .

**Theorem 12.** Consider i.i.d. alternative  $Q$  and i.i.d. null  $\mathcal{H}$  generated by  $\mathcal{P}$ . Suppose that the single-outcome reverse information projection

$$\inf_{p \in \text{conv}(\mathcal{P})} \text{KL}(q \parallel p),$$

where  $p$  ranges over the convex hull (Bayesian mixture marginals) of  $\mathcal{P}$ , is achieved by some point  $\tilde{p} \in \mathcal{P}$ . Then the likelihood ratio process

$$E(x^n) = \prod_{i=1}^n \frac{q(x_i)}{\tilde{p}(x_i)}$$

is LOAVEV for every stopping time  $\sigma$ , and it is a test super-martingale for every  $P \in \mathcal{H}$ .

Note that the strong assumption driving this strong conclusion is that the minimiser over the convex hull of  $\mathcal{P}$  happens to already be present in  $\mathcal{P}$ .

**Proof.** First let us verify that  $E$  is a test super-martingale for every  $P \in \mathcal{H}$ , from which it follows in particular that  $E$  is an anytime-valid E-value. Unit starting value  $E(\epsilon) = 1$  and non-negativity  $E \geq 0$  hold by definition. The expected value of the multiplicative increment is (recall that  $P/\tilde{P}/Q$  are i.i.d.  $p/\tilde{p}/q$ )

$$\mathbb{E}_{x \sim p} \left[ \frac{q(x)}{\tilde{p}(x)} \right] = \mathbb{E}_{x \sim q} \left[ \frac{p(x)}{\tilde{p}(x)} \right] \leq 1,$$

where the last step follows from optimality of  $\tilde{p}$  as follows. For  $\alpha \in [0, 1]$ , let  $g(\alpha) = \text{KL}(q \parallel (1 - \alpha)\tilde{p} + \alpha p)$ . As  $\tilde{p}$  is the minimiser of the reverse information projection problem while  $p$  is a feasible choice, we must have  $g'(0) \geq 0$ . We see that  $g'(0) = 1 - \mathbb{E}_q \left[ \frac{p(x)}{\tilde{p}(x)} \right]$ , proving this part of the claim.

Now let us establish that  $E$  is log-optimal. To this end, fix any other anytime-valid E-value  $Z$ . Let

$$f(\alpha) = \mathbb{E}_Q \left[ \ln((1 - \alpha)E_\sigma + \alpha Z_\sigma) \right],$$

so that it suffices to show  $f'(0) \leq 0$ . We have



$$\begin{aligned}
 f'(0) &= \mathbb{E}_Q \left[ \frac{Z_\sigma - E_\sigma}{E_\sigma} \right] \\
 &= \sum_{x^m \in \mathcal{X}^{\leq T}} Q(x^m) \left( \sigma(x^m) \prod_{i=0}^{m-1} (1 - \sigma(x^i)) \right) \frac{Z(x^m)}{E(x^m)} - 1 \\
 &\stackrel{(\star)}{=} \sum_{x^m \in \mathcal{X}^{\leq T}} \tilde{P}(x^m) \left( \sigma(x^m) \prod_{i=0}^{m-1} (1 - \sigma(x^i)) \right) Z(x^m) - 1 \\
 &= \mathbb{E}_{\tilde{P}\sigma} [Z(x^m)] - 1 \leq 0.
 \end{aligned}$$

In the first equality we differentiate and plug in  $\alpha = 0$ , while in the second we expand the definition (1) of randomised stopping time. The equality marked  $(\star)$  uses the definition of  $E$ , which gives that  $Q(x^m)/E(x^m) = \tilde{P}(x^m)$ , where  $\tilde{P}$  is i.i.d.  $\tilde{p}$ . The second inequality uses that  $Z$  is an anytime-valid E-value in particular for  $\tilde{P} \in \mathcal{H}$ .  $\square$

We conclude this section with four examples of cases where the reverse information projection falls in the null.

- When the null  $\mathcal{H}$  is generated by a convex set  $\mathcal{P}$ , then (obviously) the minimiser over the convex hull is in  $\mathcal{P}$ . This includes i.e. convex sets of i.i.d. Bernoulli distributions.
- The mean-zero Gaussians vs fixed Gaussian case from Section 4.3. Note that the null is not convex in this case, yet the reverse information projection is a point.
- $2 \times 2$  contingency tables. We take our elementary outcomes to be pairs of coin flips  $\mathcal{X} = \{0, 1\}^2$  (this corresponds to two classes with equal occurrences). Let  $p_{\theta_1, \theta_2}$  be the distribution on  $(x_1, x_2) \in \mathcal{X}^2$  claiming that  $x_1$  and  $x_2$  are independent, with  $x_1 \sim \text{Ber}(\theta_1)$  and  $x_2 \sim \text{Ber}(\theta_2)$ . In this example the null is the (i.i.d. extension of)  $\mathcal{P} = \{p_{\theta, \theta} \mid \theta \in [0, 1]\}$  and the alternative is the i.i.d. extension of  $q = p_{\theta_1, \theta_2}$  for  $\theta_1 \neq \theta_2$ . We will show that the reverse information projection is the point  $\tilde{p} = p_{\tilde{\theta}, \tilde{\theta}} \in \mathcal{P}$  where  $\tilde{\theta} = \frac{\theta_1 + \theta_2}{2}$  (as was established earlier by Turner 9, Section 2.3). Using that  $\tilde{p}$  is the reverse information projection of  $q$  onto  $\mathcal{P}$  iff  $\mathbb{E}_{x \sim p} \left[ \frac{q(x)}{\tilde{p}(x)} \right] \leq 1$  for each  $p \in \mathcal{P}$  (which we can see e.g. by reversing the steps in the first part of the proof of Theorem 12), it suffices to show that  $\mathbb{E}_{(x_1, x_2) \sim p_{\theta_1, \theta_2}} \left[ \frac{p_{\theta, \theta}(x_1, x_2)}{p_{\tilde{\theta}, \tilde{\theta}}(x_1, x_2)} \right] \leq 1$  for every  $\theta \in [0, 1]$ . Using independence and  $ab \leq (\frac{a+b}{2})^2$ , we indeed find for every  $\theta \in [0, 1]$  that

$$\begin{aligned}
 \mathbb{E}_{(x_1, x_2) \sim p_{\theta_1, \theta_2}} \left[ \frac{p_{\theta, \theta}(x_1, x_2)}{p_{\tilde{\theta}, \tilde{\theta}}(x_1, x_2)} \right] &= \mathbb{E}_{x_1 \sim \theta_1} \left[ \frac{\theta(x_1)}{\tilde{\theta}(x_1)} \right] \cdot \mathbb{E}_{x_2 \sim \theta_2} \left[ \frac{\theta(x_2)}{\tilde{\theta}(x_2)} \right] \\
 &\leq \left( \frac{1}{2} \mathbb{E}_{x_1 \sim \theta_1} \left[ \frac{\theta(x_1)}{\tilde{\theta}(x_1)} \right] + \frac{1}{2} \mathbb{E}_{x_2 \sim \theta_2} \left[ \frac{\theta(x_2)}{\tilde{\theta}(x_2)} \right] \right)^2 \\
 &= \left( \mathbb{E}_{x \sim \tilde{\theta}} \left[ \frac{\theta(x)}{\tilde{\theta}(x)} \right] \right)^2 = 1.
 \end{aligned}$$

The example readily extends to  $M \times K$  contingency tables with  $M$  classes and  $K$  outcomes. Also note that that the null is not convex in this case (the Bernoulli model for one outcome is convex, but the i.i.d. Bernoulli model for pairs of outcomes is not convex).

- Consider a one-dimensional exponential family with probability density  $p_\beta(x) = e^{\beta x - \phi(\beta)} r(x)$  compared to carrier density  $r$ , with log-partition function  $\phi(\beta) = \ln \int e^{\beta x} r(x) dx$ . Consider the one-outcome null  $\mathcal{P} = \{p_\beta \mid \beta \in [a, b]\}$  based on the parameter interval  $[a, b]$  and take the one-outcome alternative  $q = p_{\beta_1}$  for  $\beta_1 < a$  outside the null interval (the case  $\beta_1 > b$  is symmetric). We claim that  $\tilde{p} = p_a$  is the reverse information projection, and hence we have LOAVEV test super-martingale process

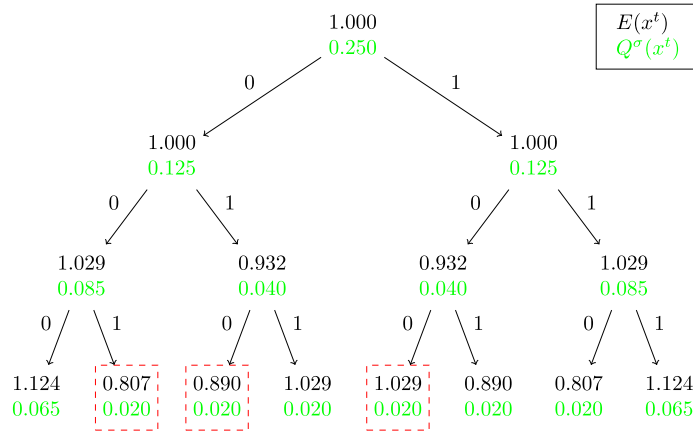
$$E(x^m) = \prod_{i=1}^m \frac{e^{\beta_1 x_i - \phi(\beta_1)} r(x_i)}{e^{a x_i - \phi(a)} r(x_i)} = e^{(\beta_1 - a) \sum_{i=1}^m x_i - m(\phi(\beta_1) - \phi(a))}.$$

Again reversing the steps in the first part of the proof of Theorem 12, it suffices to show that for every  $\beta \in [a, b]$ ,

$$\mathbb{E}_{x \sim p_{\beta_1}} \left[ \frac{p_\beta(x)}{p_a(x)} \right] \leq 1.$$

Expanding the definition of the exponential family density  $p_\beta$ , we have

$$\ln \mathbb{E}_{x \sim p_{\beta_1}} \left[ \frac{p_\beta(x)}{p_a(x)} \right] = -\phi(\beta_1) - \phi(\beta) + \phi(a) + \phi(\beta_1 + \beta - a).$$



**Fig. 2.** Numerical example with binary outcomes  $\mathcal{X} = \{0, 1\}$  and horizon  $T = 3$ . The finite sequences  $\mathcal{X}^{\leq T}$  are displayed as the nodes of a binary tree. Each node  $x^t$  is labelled with two numbers. In green, we show  $Q^\sigma(x^t)$ , i.e. the probability of seeing  $x^t$  and then stopping under the alternative  $(Q, \sigma)$ . In black, we see the LOAVEV E-value  $E(x^t)$ . The nodes marked with red dashes have the same statistics (two zeros and one one) yet they have different LOAVEV E-values.

The argument is finished by invoking convexity of  $\phi$ , which gives that  $\phi(\beta_1 + \beta - a) - \phi(\beta_1)$  is increasing in  $\beta_1$ , which is  $\leq a$ , and hence it is bounded above by  $\phi(\beta) - \phi(a)$ .

4.5. Numerical example

For this example, our aim is to go fundamentally beyond the simple LOAVEVs of the previous examples, and explore the general case numerically. To keep things tractable, we take as our null the finite set of i.i.d. Bernoulli distributions  $\mathcal{H} = \{\text{Ber}(\theta) \mid \theta \in \{0.3, 0.4, 0.5, 0.6, 0.7\}\}$ . We take as our alternative the distribution on sequences  $\{0, 1\}^T$  of coin flips formed by taking an equal parts Bayesian mixture of i.i.d. Bernoulli 0.2 and 0.8, i.e.  $Q = \frac{1}{2}\text{Ber}(0.2) + \frac{1}{2}\text{Ber}(0.8)$ . We fix a time horizon  $T$  (we will study  $T = 3$  and  $T = 6$  below), and we let the randomised stopping time  $\sigma$  stop at a uniformly random time from  $\{0, 1, \dots, T\}$  (hence the conditional probability of stopping after  $t \in \{0, \dots, T\}$  rounds is  $1/(1 + T - t)$ ). We compute the LOAVEV using numerical convex optimisation (we used the CVX add-on for MATLAB). Note that  $Q^\sigma$  puts mass on all sequences in  $\{0, 1\}^{\leq T}$ , rendering the objective strictly concave, ensuring the LOAVEV is unique. We now present the solution for  $T = 3$  and  $T = 6$ , and comment on its features.

4.5.1. Horizon  $T = 3$

For our first non-trivial case, we look at horizon  $T = 3$ . The resulting E-value is shown in Fig. 2. The implied target (value of the optimisation problem) is 0.002413. What is interesting is the following. We first argue that all ingredients of the problem only care about counts of outcomes; they do not care about the specific order of the data. Here is why: the elements of the null are i.i.d. and hence exchangeable. The alternative is a Bayesian mixture of i.i.d. distributions, and it hence is also exchangeable. The stopping time ignores the data altogether (the rule is to conditionally stop with probability  $1/(1 + T - t)$ , for an overall uniform probability  $\frac{1}{T+1}$  of stopping after  $t \in \{0, 1, \dots, T\}$  rounds). The stopping probability is hence also not making distinctions dependent on the specific order of the data. Yet the LOAVEV is order-dependent. For example, we can see in Fig. 2 that the E-values for data sequences 001, 010 and 100 are different; they are 0.807, 0.890 and 1.029.

4.5.2. Horizon  $T = 6$

The next example, displayed in Fig. 3 is at depth  $T = 6$ . We can observe the following features of the LOAVEV  $E$ .

- We still see that  $E(x^t)$  depends on the sequence of outcomes  $x^t$ , not just on the multiset. For example after 110 we have 0.814 while after 011 we have 0.923 (obtained by looking up in Fig. 3 the bit-mirrored 100).
- We also see that nodes with the same sufficient statistics may experience different multiplication factors updating the LOAVEV value over the course of another outcome. For example in context either 100 or 010 (which have the same counts), observing a subsequent 0 causes a multiplicative update to  $E$  of  $\frac{1.091}{0.923} = 1.182$  or  $\frac{0.851}{0.816} = 1.043$  respectively. These updates are different, even though the states from which they occurred had the same statistics.
- There is a set of situations where the process  $E$  is stopped. These situations are marked with green dashes in Fig. 3. It is not the case that the stopping time  $\sigma$  of the alternative stops there deterministically. This shows that the choice of stopping times in the LOAVEV can be interesting.

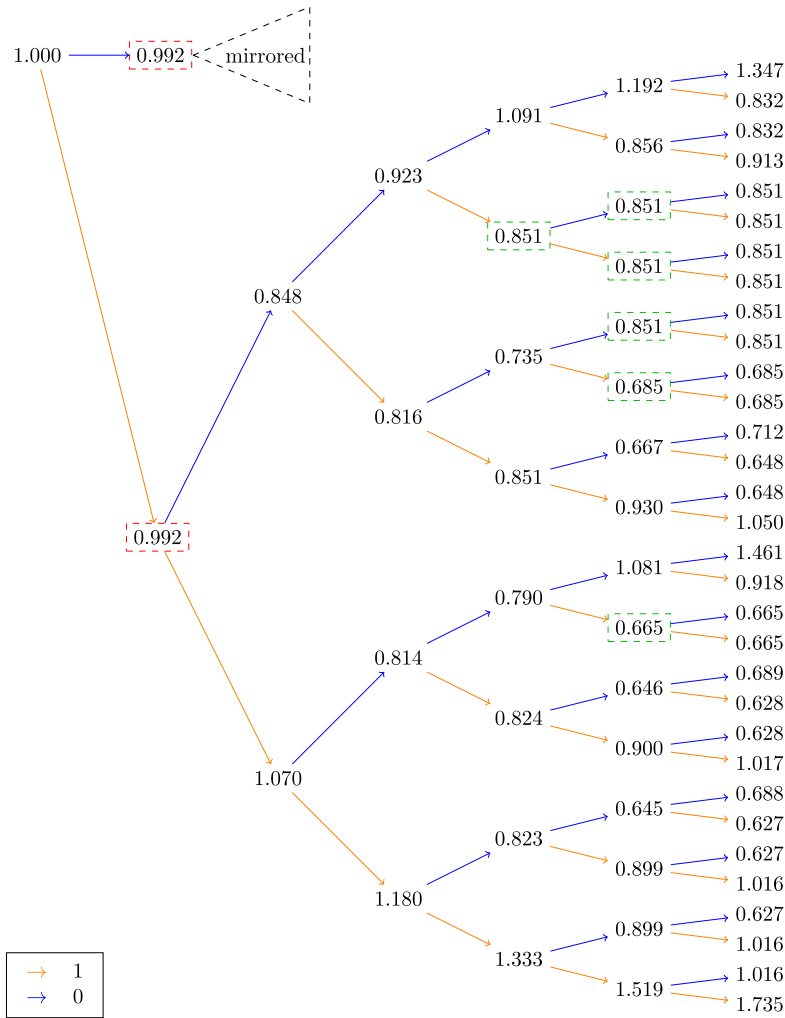


Fig. 3. LOAVEV E-value for horizon  $T = 6$ . The nodes at depth 1 (marked with red dashes) are interesting, as we have  $E(x) < 1$  for both outcomes  $x \in \{0, 1\}$ . The LOAVEV first drops deterministically, and subsequently rises back up, unlike any (super-)martingale. At the nodes marked with green dashes, the LOAVEV stops evolving before the final time  $T$ .

- Even though the null is of size  $|\mathcal{H}| = 5$ , the support of the Bayesian mixture in the dual formulation (or the active minimisers in the primal formulation) are only the two extreme points (0.3 and 0.7) of the null. Note that taking product distributions over  $T$  samples destroys convexity, even if the hypothesis set  $\mathcal{H}$  had it at the level of 1 sample.
- The implied target (i.e. the value of the LOAVEV optimisation problem) is 0.021367. For more discussion on the implied target, see [5].
- The LOAVEV  $E$  is doing something that a super-martingale cannot: it is losing evidence at first which is then later regained. To see this, observe that the two nodes at depth 1, marked with red dashes in Fig. 3, are both  $< 1$ . So the LOAVEV is losing evidence *deterministically*. To explain why this happens, we refer back to the characterisation in Lemma 4. We see that  $E$  is built from martingales (which preserve evidence on average) by taking a point-wise minimum.

### 5. Discussion

We conclude the paper with a short discussion.

- How to pick the alternative  $(Q, \sigma)$ ? In this paper we treat  $(Q, \sigma)$  as given, and optimise power (as represented by the log-optimality or implied target objective function). What about a composite alternative, represented by a set  $\mathcal{H}_1$  of candidate  $(Q, \sigma)$ ? One natural option is to somehow represent  $\mathcal{H}_1$  by a single candidate. This may be achieved e.g. by mixing with certain universal priors, as studied in the literature on universal coding. Another avenue is to take a

(stratified) worst-case approach, as is done in [2]. In principle we can maximise the minimum implied target. If the alternative  $\mathcal{H}_1$  is not separated from the null  $\mathcal{H}$ , then this trivialises. We may remove from  $\mathcal{H}_1$  the hypotheses too close to  $\mathcal{H}$ , thus creating separation, and then maximise the minimum growth rate on the remainder.

- We motivated our paper by the desire to do testing sequentially, and we obtained our log-optimal anytime-valid E-values to measure evidence against the composite null  $\mathcal{H}$  under optional stopping. Now let’s talk briefly about a sequence of such optionally stopped interactions. Length two is already interesting. That is, suppose we receive  $x^{(1)} \in \mathcal{X}^{\leq T}$ , followed by  $x^{(2)} \in \mathcal{X}^{\leq T}$ . What total evidence against the null can we report based on both? The natural measure to report is the product  $E^{\text{prod}} := E(x^{(1)})E(x^{(2)})$ . It is clear that this is safe in the sense that the expectation  $\mathbb{E}_{x^{(1)} \sim p^{\tau_1}, x^{(2)} \sim p^{\tau_2}} [E^{\text{prod}}] \leq 1$  whenever  $P \in \mathcal{H}$  and  $\tau_1, \tau_2$  are arbitrary randomised stopping times. However, something even stronger holds. Namely, for every  $P_1 \in \mathcal{H}$  and for every  $P_2 \in \mathcal{H}$  adaptively chosen based on  $x^{(1)}$ , we also have  $\mathbb{E}_{x^{(1)} \sim p_1^{\tau_1}, x^{(2)} \sim p_2^{\tau_2}} [E^{\text{prod}}] \leq 1$ . That is,  $E^{\text{prod}}$  is safe even if the true distribution is changing (possibly adversarially) in between experiments. This is the setting often studied in the literature on imprecise probability. On the one hand, we may think of robustness to a moving true distribution as a desirable safety feature. In that setting [10] study the special position product has among all e-merging functions. On the other hand, when a moving true distribution is deemed unrealistic (the classical viewpoint in statistics) we may also see this robustness as an impediment to power under the alternative. (An interesting extreme case of this phenomenon was recently studied by [4], who link it to fork-convexity.) Our Lemma 4 provides a path toward an alternative measure of evidence that removes this inefficiency. That is, we may decompose our one-experiment LOAVEV as a minimum of test martingales  $E(x^m) = \inf_{P \in \mathcal{H}} M^P(x^m)$ . This allows us to construct the sharper measure  $E^{\text{smart}} := \inf_{P \in \mathcal{H}} M^P(x^{(1)})M^P(x^{(2)}) \geq E^{\text{prod}}$ . By taking the minimum only once we lose the ability to cope with changing  $P_1, P_2 \in \mathcal{H}$ , and in return we can report the higher evidence  $E^{\text{smart}} \geq E^{\text{prod}}$  against any fixed  $P \in \mathcal{H}$ . It is an open question if any mileage can be extracted from the even stronger assumption that the subsequent experiments in addition also use the same stopping rule  $\tau_1 = \tau_2$ .
- Can we somehow simplify the problem, for example to speed up the computations? Is there any additional structure allowing us to compute things more efficiently? It is tempting to believe there should be some remnants of sufficient statistics in the LOAVEV world. A process that is a function of the sufficient statistic is described by polynomially many numbers in  $T$  (one number for every possible outcome count vector), while our current non-statistics based processes consist of exponentially many numbers (one for each prefix). However, our numerical examples show that LOAVEVs are not literally sufficient-statistic based. It would be wonderful if we can understand and exploit what can be salvaged.
- So far, we have always worked with the natural filtration  $\mathcal{F}_t = \sigma(X^t)$ . We are aware that in settings where both null and alternative are invariant under the same group, we may think of the group as a nuisance parameter. Doing this picture justice then leads to working with a reduced filtration (essentially taking the quotient by the group). This reduces the class of stopping times, and removes constraints from the definitions of test martingales and anytime-valid E-values. [2] obtain test martingales for the case where the null quotients to a point. It would be very interesting to develop the LOAVEV theory for the general case.
- In Theorem 12 we see a phenomenon, namely that the null collapses to a point  $\tilde{P}$ , upon which the dependence on the stopping time  $\sigma$  evaporates, and the LOAVEV simplifies to the likelihood ratio  $E(x^m) = Q(x^m)/\tilde{P}(x^m)$ . This happens when, from the perspective of the alternative  $Q$ , there is a single closest element  $\tilde{P}$  of the null  $\mathcal{H}$  at every stopping time. Beyond the i.i.d. case this phenomenon is not yet well understood.
- We have assumed throughout for simplicity that some final horizon  $T$  can be specified. This is clearly not a very limiting assumption, as we will eventually face the heat death of the universe. Technically, when the support of  $\sigma$  is allowed to be unbounded, we find that Proposition 2, Lemma 4 and Theorems 12 and 8 stay valid, and the update of Problem 3 remains what we want. We would need to check Theorem 5 (which, following to the reduction to batch viewpoint discussion below it, intuitively remains plausible). On the other hand Proposition 6 and Lemma 7 would have to be redone. Then again, the role of the latter two was to simplify numerical computation, which will certainly require revision to deal with unbounded horizon.

**CRedit authorship contribution statement**

**Wouter M. Koolen:** Conceptualization, Formal analysis, Methodology, Software, Writing – original draft. **Peter Grünwald:** Conceptualization, Formal analysis, Methodology, Writing – original draft.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Appendix A. Proof of Theorem 5**

We start from the left-hand side of the theorem, which is

$$\sup_{E \in \mathcal{E}_{\mathcal{H}}} \mathbb{E}_Q [\ln E_{\sigma}] = \sup_{E \in \mathcal{E}_{\mathcal{H}}} \mathbb{E}_{x^m \sim Q^{\sigma}} [\ln E(x^m)].$$

Introducing a collection of non-negative Lagrange multipliers  $\lambda(P, \tau)$ , one for each anytime validity inequality constraint  $\mathbb{E}_P [E_\tau] \leq 1$ , we find that the above problem is equal to

$$\inf_{\lambda(P, \tau) \geq 0} \sup_{\text{process } E} \mathbb{E}_{x^m \sim Q^\sigma} [\ln E(x^m)] + \int \lambda(P, \tau) (1 - \mathbb{E}_{x^m \sim P^\tau} [E(x^m)]) d(P, \tau).$$

Optimising for each  $E(x^m)$  independently, by cancelling the derivative, gives inner maximiser

$$E(x^m) = \frac{Q^\sigma(x^m)}{\int \lambda(P, \tau) P^\tau(x^m) d(P, \tau)}.$$

Plugging this in, it remains to solve

$$\inf_{\lambda(P, \tau) \geq 0} \mathbb{E}_{x^m \sim Q^\sigma} \left[ \ln \frac{Q^\sigma(x^m)}{\int \lambda(P, \tau) P^\tau(x^m) d(P, \tau)} \right] + \int \lambda(P, \tau) d(P, \tau) - 1.$$

As the final step, we reparametrise by  $\lambda(P, \tau) = \alpha w(P, \tau)$  for some prior  $w$  normalising to one and positive scale factor  $\alpha \geq 0$ . We then find that the optimal value for  $\alpha$  is  $\alpha = 1$ , where the problem becomes the right-hand side of the theorem, i.e.

$$\inf_{\text{prior } w \text{ on } (P, \tau)} \mathbb{E}_{x^m \sim Q^\sigma} \left[ \ln \frac{Q^\sigma(x^m)}{\int w(P, \tau) P^\tau(x^m) d(P, \tau)} \right].$$

### Appendix B. Proof of Lemma 7

Let  $T$  be the time horizon. Starting from the right-hand side of Lemma 4, we find that we are an infimum of test martingales. Unpacking what that means, we need to find process  $E$  and family of processes  $(M^P)_{P \in \mathcal{H}}$  optimising

$$\begin{aligned} & \max_{\substack{E \geq 0 \\ (M^P)_{P \in \mathcal{H}}}} \mathbb{E}_{x^m \sim Q^\sigma} [\ln E(x^m)] \\ & \text{subject to } \forall P \in \mathcal{H}, 0 \leq m \leq T, x^m : M^P(x^m) \geq E(x^m) \\ & \quad \forall P \in \mathcal{H}, 0 \leq m < T, x^m : M^P(x^m) \geq \mathbb{E}_P [M^P(x^{m+1}) | x^m] \\ & \quad \forall P \in \mathcal{H} : 1 \geq M^P(\epsilon). \end{aligned}$$

Introducing collections of Lagrange multipliers  $\lambda(P, x^m)$  and  $\rho(P, x^m)$  for the first two constraints, we find

$$\begin{aligned} & \min_{\lambda \geq 0, \rho \geq 0} \max_{\substack{E \geq 0, (M^P)_{P \in \mathcal{H}} \\ \forall P \in \mathcal{H} : 1 \geq M^P(\epsilon)}} \mathbb{E}_{x^m \sim Q^\sigma} [\ln E(x^m)] + \\ & \quad \sum_{0 \leq m \leq T} \sum_{P \in \mathcal{H}} \sum_{x^m} \lambda(P, x^m) (M^P(x^m) - E(x^m) +) \\ & \quad \sum_{0 \leq m < T} \sum_{P \in \mathcal{H}} \sum_{x^m} \rho(P, x^m) (M^P(x^m) - \mathbb{E}_P [M^P(x^{m+1}) | x^m]). \end{aligned}$$

We solve for  $E$  to find

$$E(x^m) = \frac{Q^\sigma(x^m)}{\sum_{P \in \mathcal{H}} \lambda(P, x^m)}$$

The coefficient on  $M^P(x^m)$  is

$$\begin{cases} \lambda(P, x^m) + \rho(P, x^m) & m = 0, \\ \lambda(P, x^m) + \rho(P, x^m) - \rho(P, x^{m-1}) P(x_m | x^{m-1}) & 0 < m < T, \\ \lambda(P, x^m) - \rho(P, x^{m-1}) P(x_m | x^{m-1}) & m = T. \end{cases}$$

The first case reveals that, at optimality,  $M^P(\epsilon) = 1$ . From the second and third cases above we further obtain the constraint that, for each length  $m > 1$ ,  $\lambda(P, x^m) \leq \rho(P, x^{m-1}) P(x_m | x^{m-1}) - \rho(P, x^m) \mathbf{1}_{m < T}$ . We are hence left with the optimisation problem

$$\min_{\lambda \geq 0, \rho \geq 0} \mathbb{E}_{x^m \sim Q^\sigma} \left[ \ln \frac{Q^\sigma(x^m)}{\sum_{P \in \mathcal{H}} \lambda(P, x^m)} \right] - 1 + \sum_{P \in \mathcal{H}} (\lambda(P, \epsilon) + \rho(P, \epsilon))$$

subject to

$$\forall P \in \mathcal{H}, 1 \leq m \leq T, x^m : \lambda(P, x^m) \leq \rho(P, x^{m-1})P(x_m | x^{m-1}) - \rho(P, x^m)\mathbf{1}_{m < T}.$$

We may further solve for  $\lambda$ . For  $1 \leq m \leq T$ , we saturate the constraint. For  $m = 0$  we find optimal value  $\sum_{P \in \mathcal{H}} \lambda(P, \epsilon) = Q^\sigma(\epsilon)$ . This results in

$$\begin{aligned} \min_{\rho \geq 0} \quad & \sum_{1 \leq m \leq T, x^m} Q^\sigma(x^m) \ln \frac{Q^\sigma(x^m)}{\sum_{P \in \mathcal{H}} (\rho(P, x^{m-1})P(x_m | x^{m-1}) - \rho(P, x^m)\mathbf{1}_{m < T})} \\ & - 1 + Q^\sigma(\epsilon) + \sum_{P \in \mathcal{H}} \rho(P, \epsilon), \end{aligned}$$

$$\text{subject to } \forall P \in \mathcal{H}, 1 \leq m \leq T, x^m : \rho(P, x^{m-1})P(x_m | x^{m-1}) \geq \rho(P, x^m)\mathbf{1}_{m < T}.$$

Finally, reparametrising by  $\rho = \alpha\phi$  with  $\alpha \geq 0$  and  $\rho > 0$  and the normalisation constraint  $\sum_{P \in \mathcal{H}} \phi(P, \epsilon) = 1 - Q^\sigma(\epsilon)$ , we obtain

$$\begin{aligned} \min_{\alpha \geq 0, \phi \geq 0} \quad & \sum_{1 \leq m \leq T, x^m} Q^\sigma(x^m) \ln \frac{Q^\sigma(x^m)}{\alpha \sum_{P \in \mathcal{H}} (\phi(P, x^{m-1})P(x_m | x^{m-1}) - \phi(P, x^m)\mathbf{1}_{m < T})} \\ & - (1 - \alpha)(1 - Q^\sigma(\epsilon)), \end{aligned}$$

$$\text{subject to } \sum_{P \in \mathcal{H}} \phi(P, \epsilon) = 1 - Q^\sigma(\epsilon),$$

$$\forall P \in \mathcal{H}, 1 \leq m \leq T, x^m : \phi(P, x^{m-1})P(x_m | x^{m-1}) \geq \phi(P, x^m)\mathbf{1}_{m < T}.$$

The optimal choice for  $\alpha$  is 1, and we obtain

$$\min_{\phi \geq 0} \quad \sum_{1 \leq m \leq T, x^m} Q^\sigma(x^m) \ln \frac{Q^\sigma(x^m)}{\sum_{P \in \mathcal{H}} (\phi(P, x^{m-1})P(x_m | x^{m-1}) - \phi(P, x^m)\mathbf{1}_{m < T})}$$

$$\text{subject to } \sum_{P \in \mathcal{H}} \phi(P, \epsilon) = 1 - Q^\sigma(\epsilon),$$

$$\forall P \in \mathcal{H}, 1 \leq m \leq T, x^m : \phi(P, x^{m-1})P(x_m | x^{m-1}) \geq \phi(P, x^m)\mathbf{1}_{m < T}.$$

## References

- [1] T. Cover, J. Thomas, *Elements of Information Theory*, Wiley-Interscience, New York, 1991.
- [2] P.D. Grünwald, R. de Heide, W.M. Koolen, Safe testing, arXiv:1906.07801, 2019.
- [3] A. Ramdas, J. Ruf, M. Larsson, W.M. Koolen, Admissible anytime-valid sequential inference must rely on nonnegative martingales, arXiv:2009.03167, 2020.
- [4] A. Ramdas, J. Ruf, M. Larsson, W.M. Koolen, Testing exchangeability: fork-convexity, supermartingales and e-processes, *Int. J. Approx. Reason.* (2021).
- [5] G. Shafer, Testing by betting: a strategy for statistical and scientific communication, *J. R. Stat. Soc., A* 184 (2) (2020) 407–431.
- [6] G. Shafer, A. Shen, N. Vereshchagin, V. Vovk, Test martingales, Bayes factors and p-values, *Stat. Sci.* 26 (2011) 84–101.
- [7] G. Shafer, V. Vovk, *Game-Theoretic Foundations for Probability and Finance*, Wiley Series in Probability and Statistics, Wiley, Hoboken, NJ, 2019.
- [8] E. Solan, B. Tsirelson, N. Vieille, Random stopping times in stopping problems and stopping games, arXiv:1211.5802, 2012.
- [9] R.J. Turner, Safe tests for  $2 \times 2$  contingency tables and the Cochran-Mantel-Haenszel test, Master's thesis, Leiden University, 2019.
- [10] V. Vovk, R. Wang, E-values: calibration, combination and applications, *Ann. Stat.* 49 (2021) 1736–1754.