Accuracy and Decision Time for Decentralized Implementations of the Sequential Probability Ratio Test

Sandra Hala Dandach   Ruggero Carli   Francesco Bullo

Abstract— This paper focuses on decentralized decision making in a population of individuals each implementing the sequential probability ratio test. The individual decisions are combined by a fusion center into a collective decision via an aggregation rule. For distinct aggregation rules, we study how the population size affects the performance of the collective decision making, i.e., the decision accuracy and time.

We analyze two aggregation rules, the fastest rule and the majority rule. In the fastest rule, the group decision is equal to the first decision made by any individual. In the majority rule, the group decision is equal to the majority of the decisions. Under the assumption of measurement independence among individuals, we introduce a novel numerical representation of the performance of decentralized decision making. We study our settings analytically as well as numerically. Our numerical results and simulations characterize the tradeoff between accuracy and decision time as a function of the population size.

I. INTRODUCTION

This work aims to understand how grouping individual decision makers (DM) affects the speed and accuracy with which a collective decision is made. This class of problems has a rich history and some of its variations are studied in the context of distributed detection in sensor networks [1], [2], [3], [4], [5] and Bayesian learning in social networks [6], [7].

In this paper we consider a group of N individuals each of them individually implementing the standard sequential probability ratio test (SPRT) with the purpose of deciding between two hypothesis H_0 and H_1. We refer to the individuals as *decision makers* which we denote as DMs hereafter. In our setup no-communication is allowed between the DMs. Once a DM has provided a decision it communicates it to a fusion center. The fusion center collects the various decisions and provides a global decision via an aggregation rule.

In this paper we focus on two aggregation rules, the *fastest* rule and the *majority* rule. In the fastest rule, the fusion center makes a global decision as soon as one hypothesis gets more votes than the opponent hypothesis, while in the majority rule the fusion rule makes a global decision as soon as one of the two hypothesis gets $\lfloor N/2 \rfloor + 1$ decisions in its favor. For both rules and for distinct sizes of the group of DMs, we study the performance of the collective decision making, i.e., the decision accuracy and expected number of observations required to provide a decision.

The framework we analyze in this paper is related to the one considered in many papers in the literature, see for instance [8], [1], [9], [2], [10], [3], [11] and references therein. The focuses of these works are mainly two-fold. First, determining which type of information should the DMs send to the fusion center. Second, computing optimal decision rules both for the DMs and the fusion center where optimality refers to maximizing accuracy. One key implicit assumption made in these works, is the assumption that the aggregation rule is applied by the fusion center only after all the DMS have provided their local decisions.

The work presented in this paper, relaxes the assumption on the local decisions. Indeed the fusion center might provide the global decision much earlier than the time needed for the local decisions to be made by the whole group. Our main concern is the exact computation of both the accuracy of the final group decision as well as the expected number of observations required by the group, in order to provide the final decision. In this work we accomplish these objectives by proposing exact expressions for the conditional probabilities of the group giving the correct and wrong final decisions, at any time instant. We perform this analysis both for the fastest rule and the majority rule for a varying group sizes. This represents the major contribution of this paper.

In the second part of the paper we use the expressions that we provide in the first part of the paper, to numerically characterize the tradeoff between the accuracy and the expected time and the group size. For illustration, we consider a discrete distribution of the Koopman-Darmoi-Pitman form. We find that the majority rule provides, as N increases, a remarkable improvement in terms of of the accuracy while the fastest rule provides a remarkable improvement in terms of the expected number of observations required to provide the final decision.

The rest of the paper is organized as follows. In Section II we review the standard SPRT. In Section III we formally introduce the problem studied in this paper. In Section IV we present our novel numerical method useful to analyze the problem of interest. In Section V we provide some numerical results. We conclude in Section VI.

II. A BRIEF REVIEW OF SPRT AND OF RELATED ANALYSIS METHODS

In this section we discuss the classical *sequential probability ratio test* (SPRT) for a single decision maker; to do so we follow the treatment in [12].

A. Classical Wald's analysis

SPRT algorithm has been developed in 1943 by Abraham Wald for the purpose of deciding between two simple hypotheses. Suppose that the random variable X has a distribution $f(x; \theta)$ and that we want to test the null hypothesis $H_0 : \theta = \theta_0$ against the alternative hypothesis $H_1 : \theta = \theta_1$. Suppose the observations $x(1), x(2), x(3), \ldots$, are taken one at a time until a decision is made and assume that they are independent, i.e., $\mathbb{E}[x(i)x(j)] = \mathbb{E}[x(i)]\mathbb{E}[x(j)]$ if $i \neq j$. We define the log-likelihood ratio associated to k-th observation $x(k)$ as

$$\lambda(k) = \log \frac{f(x(k), \theta_1)}{f(x(k), \theta_0)}.$$

The sum of the log-likelihoods up to k is denoted by $\Lambda(k) = \sum_{j=1}^{k} \lambda(j)$. For the SPRT, sampling continues as long as $\eta_0 < \Lambda(k) < \eta_1$, where η_0 and η_1 are two pre-assigned thresholds. Sampling is stopped the first time this inequality is violated. If $\Lambda(k) < \eta_0$, then the group decides for $\theta = \theta_0$. If $\Lambda(k) > \eta_1$, then the group decides for $\theta = \theta_1$. It can be shown that, for any values of η_0 and η_1, the SPRT algorithm provides a decision almost surely.

The goal is to set the thresholds η_0 and η_1 in such a way that the probabilities of misdetection (saying H_0 when H_1 is correct, i.e., $\mathbb{P}[\text{say}H_0|H_1]$) and of false alarm (saying H_1 when H_0 is correct, i.e., $\mathbb{P}[\text{say}H_1|H_0]$) are equal to some pre-assigned values p_m and p_f (the subscripts m and f refer, respectively, to misdetection and false-alarm). Wald proved that in order to guarantee the following probabilistic constraints

$$\mathbb{P}[\text{say}H_0|H_1] = p_m \quad \text{and} \quad \mathbb{P}[\text{say}H_1|H_0] = p_f, \quad (1)$$

the values of η_0 and η_1 must satisfy the following conditions: $\eta_0 \geq \log \frac{p_m}{1-p_f}$ and $\eta_1 \leq \log \frac{1-p_m}{p_f}$. Clearly, using equalities in the above inequalities, that is,

$$\eta_0 = \log \frac{p_m}{1 - p_f} \quad \text{and} \quad \eta_1 = \log \frac{1 - p_m}{p_f}, \quad (2)$$

one obtain that $\mathbb{P}[\text{say}H_0|H_1] \leq p_m$ and $\mathbb{P}[\text{say}H_1|H_0] \leq p_f$. The classical Wald's analysis computes the values for η_0 and η_1 according to the formulas given in (2).

In [12] Wald determined also the average number of samples, denoted by T, required by the algorithm for providing a decision to be approximately

$$\mathbb{E}[T|H_0] \approx \frac{(1 - p_f)\eta_0 + p_f \eta_1}{\mathbb{E}\left[\log \frac{f(x, \theta_1)}{f(x, \theta_0)}|H_0\right]} \quad (3)$$

and

$$\mathbb{E}[T|H_1] \approx \frac{(1 - p_m)\eta_1 + p_m \eta_0}{\mathbb{E}\left[\log \frac{f(x, \theta_1)}{f(x, \theta_0)}|H_1\right]}. \quad (4)$$

The denominators that show in equations (3) and (4) are the Kullback-Leibler distances for the distributions. Observe that equations (1), (3) and (4) are exact if, upon stopping with k observations, either $\Lambda(k) = \eta_0$ or $\Lambda(k) = \eta_1$. However some overshooting of the boundaries generally occurs. The distribution of the overshooting $\Lambda(k) - \eta_1$ and $\eta_0 - \Lambda_0$,

quantify the accuracy of the approximations introduced by the formulas in (2).

B. Computation of accuracy and decision time for pre-assigned thresholds η_0 and η_1

In the previous subsection we have seen how the classical Wald's analysis provides simple formulas for setting the values of η_0, η_1 in order to meet pre-defined requirements on the probabilities of error. In this subsection we would like to address the reversed problem: Given a-priori thresholds η_0 and η_1, how can we compute the values of the $\mathbb{P}[\text{say}H_1|H_0]$, $\mathbb{P}[\text{say}H_0|H_1]$, $\mathbb{E}[T|H_0]$ and $\mathbb{E}[T|H_1]$?

We start by observing that an exact computations of those probabilities and expected values would be possible if we knew the probabilities of crossing the boundaries at any time $k \in \mathbb{N}^1$. Indeed in such case, if $p_{i|j}(k)$ denotes the probability of deciding H_i at time k, given that H_j is the correct hypothesis, i.e., $p_{i|j}(k) := \mathbb{P}[\text{say}H_i|H_j, T = k]$, from standard probabilistic arguments it follows that

$$\mathbb{P}[\text{say}H_1|H_0] = \sum_{t=1}^{\infty} p_{1|0}(t), \quad (5)$$

and

$$\mathbb{E}[T|H_0] = \sum_{t=1}^{\infty} (p_{0|0}(t) + p_{1|0}(t))t. \quad (6)$$

Similar formulas apply for $\mathbb{P}[\text{say}H_0|H_1]$ and $\mathbb{E}[T|H_1]$.

The author in [10], provided a recursive method for computing exact values of these probabilities which can be applied to a broad class of discrete distributions, precisely when X is a discrete random variable of the Koopman-Darmoi-Pitman form. The approach we will take later in analyzing a group of DMs assumes the knowledge of the quantities $\{p_{i|j}(t)\}_{t=1}^{\infty}$, $i, j = 1, 2$. We proceed now by reviewing the procedure proposed in [10] for computing these terms.

Let X be a discrete random variable of the Koopman-Darmoi-Pitman form; that is

$$f(x, \theta) = \begin{cases} h(x) \exp(B(\theta)Z(x) - A(\theta)) & \text{if} \quad x \in \mathcal{Z} \\ 0 & \text{if} \quad x \notin \mathcal{Z}, \end{cases}$$

where $h(x)$, $Z(x)$ and $A(\theta)$ are known functions and where $\mathcal{Z}$ is a subset of the integer numbers $\mathbb{Z}$. In this section we shall assume that $Z(x) = x$. It is worth noting that the Bernoulli, binomial, geometric, negative binomial and Poisson distributions are some widely used distributions of the Koopman-Darmoi-Pitman form, satisfying the condition $Z(x) = x$. For distributions of this form we have that

$$\lambda(k) = (B(\theta_1) - B(\theta_0))x(k) - (A(\theta_1) - A(\theta_0)).$$

One can easily see that sampling will continue as long as

$$\frac{\eta_0 + k(A(\theta_1) - A(\theta_0))}{B(\theta_1) - B(\theta_0))} < \sum_{i=1}^{k} x(i) < \frac{\eta_1 + k(A(\theta_1) - A(\theta_0))}{B(\theta_1) - B(\theta_0))}$$

---

[1]By convention $\mathbb{N}$ denotes the set of positive integer numbers, i.e., $\mathbb{N} = \{1, 2, 3, \ldots\}$

for $B(\theta_1) - B(\theta_0) > 0$; if $B(\theta_1) - B(\theta_0) < 0$ the inequalities would be reversed. Observe that $\sum_{i=1}^{k} x(i)$ is an integer number. Now let $\bar{\eta}_0^{(k)}$ be the smallest integer greater than $\{\eta_0 + k(A(\theta_1) - A(\theta_0))\}/(B(\theta_1) - B(\theta_0))$ and let $\bar{\eta}_1^{(k)}$ be the largest integer smaller than $\{\eta_1 + k(A(\theta_1) - A(\theta_0))\}/(B(\theta_1) - B(\theta_0))$. Sampling will continue as long as $\bar{\eta}_0^{(k)} \leq \mathcal{X}(k) \leq \bar{\eta}_1^{(k)}$ where $\mathcal{X}(k) = \sum_{i=1}^{k} x(i)$. Now suppose that, for any $\ell \in [\bar{\eta}_0^{(k)}, \bar{\eta}_1^{(k)}]$ the probability $\mathbb{P}[\mathcal{X}(k) = \ell]$ is known. Then we have

$$\mathbb{P}[\mathcal{X}(k+1) = \ell | H_i] = \sum_{j=\bar{\eta}_0^{(k)}}^{\bar{\eta}_1^{(k)}} f(\ell - j; \theta_i) \mathbb{P}[\mathcal{X}(k) = j | H_i],$$
(7)

and

$$p_{1|i}(k+1) = \sum_{j=\bar{\eta}_0^{(k)}}^{\bar{\eta}_1^{(k)}} \sum_{r=\bar{\eta}_1^{(k)}-j+1}^{\infty} \mathbb{P}[\mathcal{X}(k) = j | H_i] f(r; \theta_i)$$

$$p_{0|i}(k+1) = \sum_{j=\bar{\eta}_0^{(k)}}^{\bar{\eta}_1^{(k)}} \sum_{r=-\infty}^{\bar{\eta}_0^{(k)}-j-1} \mathbb{P}[\mathcal{X}(k) = j | H_i] f(r; \theta_i). \quad (8)$$

Starting with $\mathbb{P}[\mathcal{X}(0) = 1]$ it is possible to compute recursively all the quantities $\{p_{i|j}(k)\}_{k=1}^{\infty}$ and $\mathbb{P}[\mathcal{X}(k) = \ell]$, for any $k \in \mathbb{N}$, $\ell \in [\bar{\eta}_0^{(k)}, \bar{\eta}_1^{(k)}]$, and $\{p_{i|j}(t)\}_{t=1}^{\infty}$.

III. DECENTRALIZED SPRT: PROBLEM SETUP

In this Section we present our problem of interest. Consider a group of N decision makers which are repeatedly observing a random variable X of distribution $f(x, \theta)$, with the purpose of choosing between the two hypothesis $H_0 : \theta = \theta_0$ and $H_1 : \theta = \theta_1$. Suppose each decision maker is individually implementing the classic SPRT, reviewed in Section II-A, with the assumptions that

(i) all the decision makers have the same thresholds η_0 and η_1;

(ii) no-communication is allowed between the decision makers;

(iii) the probability distributions $f(x, \theta_0)$ and $f(x, \theta_1)$ are known a-priori; and

(iv) the observations taken, conditioned on either hypothesis, are independent in time as well as from decision maker to decision maker.

Once a decision maker arrives at a final local decision, it communicates it to a fusion center. The fusion center collects the messages it receives keeping track of the number of decisions in favor of H_0 and in favor of H_1. A global decision is provided by the fusion center according to one of the following two stopping rules introduced in the literature [1] : the *fastest* rule and the *majority* rule. Informally in the *fastest* rule the decision of the group is set equal to the decision of the fastest decision maker while in the *majority* rule the group's decision is set equal to the agreed on decision by the majority of votes. In what follows, we refer to the framework described above as to the *Decentralized SPRT with the **fastest/majority*** rule.

Next we formally describe our setup. Let N denote the size of the group of decision makers. Then the *Decentralized SPRT with the* fastest/majority *rule* is defined as follows:

Processor states : For each $i \in \{1, \ldots, N\}$, the i-th DM stores in memory the variable Λ_i and two pre-assigned values for the thresholds η_0 and η_1.

Initialization : All variables Λ_i, $i \in \{1, \ldots, N\}$, are initialized to 0, i.e., $\Lambda_i(0) = 0$, $i \in \{1, \ldots, N\}$.

State iteration (Decision Makers) : At time $k \in \mathbb{N}$, for each i, if $\Lambda_i(k-1) \in [\eta_0, \eta_1]$, the i-th DM performs three actions in the following order:

(1) DM i takes a new measurement $x_i(t)$ and computes the log-likelihood

$$\lambda_i(k) = \log \frac{f(x_i(k), \theta_1)}{f(x_i(k), \theta_0)}. \quad (9)$$

(2) DM i updates the variable Λ_i by

$$\Lambda_i(k) = \Lambda_i(k-1) + \lambda_i(k). \quad (10)$$

(3) DM i compares $\Lambda_i(k)$ with the thresholds η_0 and η_1. In case either $\Lambda_i(k) < \eta_0$ or $\Lambda_i(k) > \eta_1$, DM i sends a message u_i to the fusion center to communicate its local decision; specifically

$$u_i = \begin{cases} -1 & \text{if } \Lambda_i(k) < \eta_0 \\ +1 & \text{if } \Lambda_i(k) > \eta_1. \end{cases}$$

After having send its message u_i the i-th decision maker ceases running its SPRT algorithm.

The last step is to introduce the stopping rules that the fusion center adopts to provide a decision. Below we formally describe the *fastest* rule and *majority* rule.

fastest rule : Define the following two sets

$$S_+(k) = \{i \in V : \Lambda_i(k) \geq \eta_1\}$$

and

$$S_-(k) = \{i \in V : \Lambda_i(k) \leq \eta_0\}$$

and the corresponding variables

$$Count_+(k) = |S_+(k)|$$

and

$$Count_-(k) = |S_-(k)|$$

Let T be the random variable accounting for the time when the fusion center provides its decision, then

$$T = \inf \{k \in \mathbb{N} : Count_+(k) - Count_-(k) \neq 0\}.$$

If $Count_+(T) > Count_-(T)$ then the fusion center decides in favor of H_1; otherwise if $Count_-(T) > Count_+(T)$ then the fusion center decides in favor of H_0.

majority rule: Let $S_+, S_-, Count_+, Count_-, T$ be defined as in the previous rule, then

$$T = \inf \left\{ k \in \mathbb{N} : \text{either } Count_+(k) \geq \left\lceil \frac{N}{2} \right\rceil \right.$$
$$\left. \text{or } Count_-(k) \geq \left\lceil \frac{N}{2} \right\rceil \right\}.$$

If $Count_+(T) \geq \lceil \frac{N}{2} \rceil$ then the fusion center decides in favor of H_1; otherwise if $Count_-(T) \geq \lceil \frac{N}{2} \rceil$ the fusion center decides in favor of H_0.

Given a group of decision makers, the goal is to compute accuracy and decision time, both for the *Decentralized SPRT with the* fastest *rule* and the *Decentralized SPRT with the* majority *rule*. We will focus on this problem in the next Section.

We conclude this Section by stating a Proposition that establishes a desirable stopping property for both the above stopping rules, when N is odd.

Proposition III.1 *Let $N \in \mathbb{N}$ be odd and consider the* Decentralized SPRT with the *fastest/majority* rule *described above. Then, both the* fastest *stopping rule and the* majority *stopping rule, reach a decision almost surely.*

We omit the straightforward proof but we draw the attention of the reader that the proof of this statement follows with minor changes from Wald's proof of almost sure convergence for SPRT. In the rest of the paper we will assume the following property.

Assumption III.1 *The size of the group of decision makers, i.e., N, is odd.*

IV. DECENTRALIZED SPRT: ACCURACY AND DECISION TIME OF THE FASTEST AND MAJORITY RULES

The goal of this Section is to analyze accuracy and expected decision time for the Decentralized SPRT both with the *fastest* rule and the *majority* rule. The basic idea is to provide a recursive method to compute, for a group of DMs, the exact probabilities of providing a decision at any time $t \in \mathbb{N}$. Similarly to Equations (5), (6) these probabilities will allow to estimate the accuracy and the expected decision time.

In what follows, we denote by $p_{i|j}(t, N)$ the probability that, given a group of N decision makers, the fusion center provides its global decision at time t in favor of hypothesis H_i, given that H_j is correct, i.e.,

$p_{i|j}(t, N) =$
$\quad \mathbb{P}\left[\text{Global decision for } N \text{ DMs is } H_i | H_j, T = t\right].$

It is worth observing that $p_{i|j}(t, 1)$ coincides with the probability $p_{i|j}(t)$ defined in Section II-B when one DM is running SPRT. For ease of notation will continue using $p_{i|j}(t)$ instead of $p_{i|j}(t, 1)$.

In what follows we provide formulas to compute recursively the probabilities $\{p_{i|j}(t, N)\}_{t=1}^{\infty}$, $i, j = 0, 1$. We carry on under the assumption that the quantities $\{p_{i|j}(t)\}$, i.e., the probabilities for the standard SPRT with one DM, are known a-priori. In particular we can assume that they have been computed by the procedure described in Section II-B.

A. Decentralized SPRT with the fastet rule

We start our analysis by introducing three useful quantities.

First, define the probability function of balanced decisions, $\alpha : \mathbb{N} \times \{1, \ldots, \lfloor N/2 \rfloor\} \to [0, 1]$, as follows: $\alpha(t, k)$ is the probability that, given a group of $2k$ decision makers, there exists a sequence $k_1 \leq k_2 \leq \ldots \leq k_t = k$ such that for any $\tau \in [1, t]$, $Count_+(\tau) = Count_-(\tau) = k_\tau$. By convention we set $\alpha(t, 0) = 1$ for all $t \geq 0$ and $\alpha(0, k) = 0$ for all $k \in \{1, \ldots, \lfloor \frac{N}{2} \rfloor\}$.

Second, define the probability function of decision H_1 at time t, $\beta_{1|j} : \mathbb{N} \times \{1, \ldots, N-1\} \to [0, 1]$, $j = 0, 1$, as follows: $\beta_{1|j}(t, h)$ is the probability that, for all times $\tau \in [1, t-1]$, no decision maker in a group of size h have decided, and that at time t the number of decisions favoring hypothesis H_1 is larger than the number of decisions favoring H_0 when H_j is true. Clearly, in an analogous way we can provide also the definition of probability function of decision H_0 at time t, $\beta_{0|j} : \mathbb{N} \times \{1, \ldots, N-1\} \to [0, 1]$, $j = 0, 1$.

Finally, define the probability function of decision up to time t, $\gamma_{i|j} : \mathbb{N} \to [0, 1]$, as follows: $\gamma_{i|j}(t)$ is the probability that a single DM provided the decision H_i when H_j is correct, before or at time t, i.e.,

$$\gamma_{i|j}(t) = \sum_{s=1}^{t} p_{i|j}(t). \tag{11}$$

Proposition IV.1 *Given a group of N DMs satisfying Assumption III.1, consider the Decentralized SPRT with the* fastest *rule as described in Section III. Without loss of generality let us assume that H_1 is the correct hypothesis. Then the following statements hold.*

First, the probability that the group gives the correct decision at time $t \in \mathbb{N}$ is

$$p_{1|1}(t, N) = \sum_{s=0}^{\lfloor \frac{N}{2} \rfloor} \binom{N}{2s} \alpha(t-1, s) \beta_{1|1}(t, N-2s). \tag{12}$$

Second, the probability of balanced decisions up to time $t \in \mathbb{N}$ satisfies the recursion equations:

$$\alpha(t, k) = \sum_{s=0}^{k-1} \binom{2k}{2k-2s} \alpha(t-1, k-s) \binom{2s}{s} p_{1|1}^s(t) p_{0|1}^s(t)$$
$$+ \binom{2k}{k} p_{1|1}^k(t) p_{0|1}^k(t). \tag{13}$$

Third, the probability $\beta_{1|1}(t, h)$ is given by

$$\beta_{1|1}(t, h) = \sum_{j=1}^{\lfloor \frac{h}{2} \rfloor} \binom{h}{j} p_{11}^j(t) \times$$
$$\times \left(\sum_{i=0}^{j-1} \binom{h-j}{i} p_{0|1}^i(t) \left[1 - \gamma_{1|1}(t) - \gamma_{0|1}(t)\right]^{h-j-i} \right)$$
$$+ \sum_{j=\lceil \frac{h}{2} \rceil}^{h} \binom{h}{j} p_{1|1}^j(t) \left[1 - \gamma_{1|1}(t) - \gamma_{0|1}(t-1)\right]^{h-j}. \tag{14}$$

Proof: For reasons of space we only sketch the proof of the Proposition.

Assume H_1 is correct. When analyzing the *fastest* aggregation rule, we need to keep in mind that the time at which a DM gives its vote matters, since the group decides as soon as $Count_+(k) - Count_-(k) > 0$. Once the group decides, the later votes do not change the decision. To keep track of this, we define the following two events:

Event 1 Given a group of $h = N - 2k$ DMs, then

- $Count_+(\tau) = Count_-(\tau) = 0$, for $1 \leq \tau \leq t - 1$.
- $Count_+(t) > Count_-(t)$

$\beta_{1|1}(t, h)$ denotes the probability of event 1. Event 1 denotes the event in which, no decision has been made by the h individuals for all times up to $t - 1$, then exactly at t one of the two hypothesis is favored over the other. Event 1 happens in one of two ways. Either, $j \leq \lfloor \frac{h}{2} \rfloor$ individuals decide in favor of H_1 at time t while at most $j - 1$ individuals cast their votes in favor of H_0 at the same time (while the rest of the h individuals do not decide), or at time t more than half of the h individuals decide in favor of H_1 (while the rest either do not decide or decide for H_0 at time t). Equation (14) follows from these considerations.

Event 2 Given $2k$ DMs and $t \in \mathbb{N}$, there exists a sequence $k_1 \leq k_2 \leq \ldots \leq k_t = k$ such that for any $\tau \in [1, t]$ $Count_+(\tau) = Count_-(\tau) = k_\tau$.

$\alpha(t, k)$ denotes the probability of event 2. Event 2 denotes the event in which all decisions are made by pairs of DMs whose number equally divides between the two hypothesis, and therefore cancel out. For event 2 to happen at time t with $2k$ DM, it has to be that at time $t - 1$, $2k - 2j$ DM have canceled out and that at time t, the remaining $2j$ DMs cast their votes and cancel out, or another possibility is that all $2k$ DM decide and cancel out at time t. Equation (13) follows.

Since the DMs take independent measurements, Events 1 and 2 are independent events, equation (12) follows.

∎

Observe that in the previous Proposition, $\beta_{1|1}(t, h)$ can be computed offline for all t and h. On the other hand, $\alpha(t, k)$ is computed recursively; in order to calculate $\alpha(t, k)$, we need to have $\alpha(t - 1, j)$ for all $j \in \{1, 2, \ldots, N - 1\}$.

B. Decentralized SPRT with the majority rule

Similar to the previous method, we present in this part, the iterative numerical method through which we are able to compute the probability of the fusion center making a decision, as a function of the exact probabilities presented in section II and of the size of the group, when the group applies the majority rule.

Proposition IV.2 *Consider the Decentralized SPRT with the* majority *rule described above for a group of size N of decision makers under Assumption III.1. Then for any $t \in \mathbb{N}$*

we have that

$$p_{i|j}(t, N) = \sum_{k=0}^{\lceil \frac{N}{2} \rceil - 1} \sum_{j=\lceil \frac{N}{2} \rceil - k}^{N-k} \binom{N}{k} \gamma_{i|j}^k(t-1) \times$$
$$\binom{N-k}{j} p_{i|j}^j(t) \left(1 - \gamma_{i|j}(t)\right)^{N-(j+k)},$$

where $\gamma_{i|j}(t)$ was defined in (11).

The proof of this proposition follows the same line of thought as that of proposition IV.1. We omit the proof for the lack of space, but we briefly state the intuition. The *majority* rule will decide at time t in favor H_i when H_j is correct, if and only if, at least $\lceil \frac{N}{2} \rceil$ DMs vote for H_i, while at most $\lfloor \frac{N}{2} \rfloor$ DMs voted for H_j, at time t. It is not necessary that all the DMs would have voted before the group decides, leaving room for DMs who do not cast their opinions.

V. NUMERICAL RESULTS

In this section we provide an example of the numerical method we describe in the previous Section. Both for the *fastest* rule and the *majority* rule, given a group of N decision makers, we compute the probabilities $\{p_{i|j}(t, N)\}$, $i, j \in \{0, 1\}$; thanks to these quantities we estimate accuracy and expected decision time according to the formulas

$$\mathbb{P}[\text{say} H_i | H_j] = \sum_{t=1}^{\infty} p_{i|j}(t, N),$$

and

$$\mathbb{E}[T | H_i] = \sum_{t=1}^{\infty} (p_{0|i}(t, N) + p_{1|i}(t, N))t.$$

In this Section we assume that X is a distribution of the Koopman-Darmoi-Pitman form and hence that, through the procedure illustrated in Section II, we can have access, at all times t, to the probability $p_{ij}(t)$. Specifically let $f(x; \theta)$ be given by the binomial distribution, i.e.,

$$f(x; \theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}$$

for $x \in \{0, 1, \ldots, n\}$, and $f(x; \theta) = 0$ otherwise. We assume here that $n = 5$.

In this Section we made our computations for three different pairs (θ_0, θ_1), under the assumption that H_1 is always the correct hypothesis. Precisely we consider $\theta_0 = 0.5 - \epsilon$ and $\theta_1 = 0.5 + \epsilon$ where $\epsilon = \{0.02; 0.05; 0.08\}$. Moreover, in order to point out the tradeoff accuracy-expected decision time vs number of decision makers, we analyze group of increasing size.

In all the scenarios we consider, we use the same values for the thresholds η_0 and η_1. In particular we assumed that η_0 and η_1 are computed according to the formulas (2) letting $p_m = p_f = 0.1$.

The results obtained are reported in Figures 1, 2 and 3. We can see that, as the size of the group increases, the majority rule provides a remarkable improvement in terms of the accuracy, while the fastest rule provides a remarkable

improvement in terms of the expected number of iterations required to provide a decision. The fastest rule improves also the accuracy even though this improvement is not as significant as the one provided by the majority rule. The majority rule, in terms of number of iterations to provide a decision, exhibits performance which are almost constant as N varies.

Moreover, observe that the bigger the value of ϵ the faster and the more accurate the decision making scheme is. This is due to the fact that, as ϵ increases, the Kullback-Leibler distance between the distributions $f(x, \theta_0)$ and $f(x, \theta_1)$ increases.
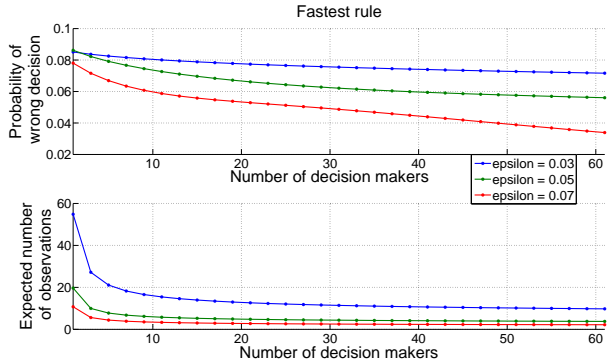

Fig. 1. In the first subplot we show the probability of error of the fastest decision rule as the number of DMs changes. In the second subplot we show the expected number of observations needed for the same group of DMs applying the same rule to make a decisions.
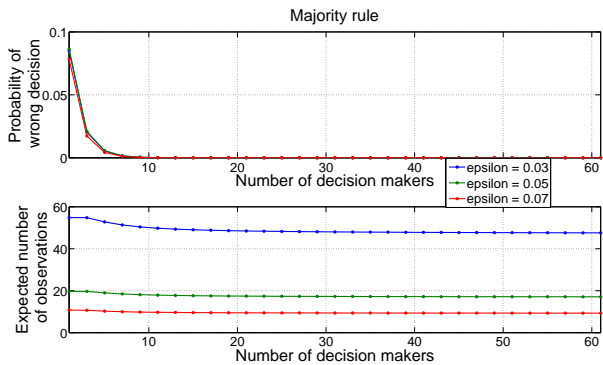

Fig. 2. In the first subplot we show the probability of error of the majority rule as the number of DMs changes. In the second subplot we show the expected number of observations needed for the same group of DMs applying the same rule to make a decisions.

VI. CONCLUSION

In this work, we presented a complete analysis of how a group of DMs can collectively reach a decision about the correctness of a hypothesis. We presented a numerical method that made it possible to completely analyze and understand interesting fusion rules of the individuals decisions. The analysis we presented concentrated on two aggregation rules, the fastest rule and the majority rule. We provided a numerical example showing that, as the size of the group
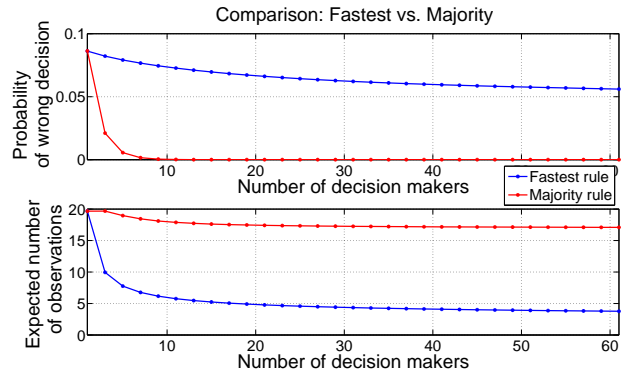

Fig. 3. In the first subplot we show a comparison between the probability of error of the majority and the fastest rule as the number of DMs changes. In the second subplot we show the expected number of observations needed for the DMs applying the same rules to make a decisions

of decision makers increases, the majority rule significantly improves the accuracy of the decision scheme, while the fastest rule significantly reduces the expected number of iterations required to provide a decision. It will be the subject of our future work to estimate these improvements as function of the number of DMs. Moreover we aim to extend our analysis to the case where the decision makers are allowed to communicate with each other.

REFERENCES

[1] P. K. Varshney, *Distributed Detection and Data Fusion*. Signal Processing and Data Fusion, Springer, 1996.

[2] J. N. Tsitsiklis, *Problems in Decentralized Decision Making and Computation*. PhD thesis, Massachusetts Institute of Technology, Nov. 1984. Available at http://web.mit.edu/jnt/www/Papers/PhD-84-jnt.pdf.

[3] V. V. Veeravalli, T. Başar, and H. V. Poor, "Decentralized sequential detection with sensors performing sequential tests," *Mathematics of Control, Signals and Systems*, vol. 7, no. 4, pp. 292–305, 1994.

[4] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407–416, 2003.

[5] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE Journal on Selected Areas in Communication*, vol. 22, pp. 1007–1015, Aug. 2004.

[6] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," Working Paper 14040, National Bureau of Economic Research, May 2008.

[7] A. Tahbaz-Salehi, A. Sandroni, and A. Jadbabaie, "Preliminary results on learning under social influence," in *IEEE Conf. on Decision and Control*, (Shanghai, China), pp. 1513–1519, Dec. 2009.

[8] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, vol. 2, pp. 297–344, 1993.

[9] W. W. Irving and J. N. Tsitsiklis, "Some properties of optimal thresholds in decentralized detection," *IEEE Transactions on Automatic Control*, vol. 39, no. 4, pp. 835–838, 1994.

[10] L. J. Young, "Computation of some exact properties of Wald's SPRT when sampling from a class of discrete distribution," *Biometrical Journal*, vol. 36, no. 5, pp. 627–637, 1994.

[11] V. V. Veeravalli, "Decentralized quickest change detection," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1657–1665, 2001.

[12] A. Wald, "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, 1945.