

Accuracy and Decision Time for a Class of Sequential Decision Aggregation Rules

Sandra Hala Dandach

Ruggero Carli

Francesco Bullo

Abstract—This work focuses on decentralized decision making in a population of individuals each implementing the sequential probability ratio test. The individual decisions are combined into a decentralized decision via an aggregation rule chosen from a family of aggregation rules, denoted as *q out of N* rule. We study how the population size affects the performance of the decentralized decision making, i.e., the decision accuracy and time. In a group applying the *q out of N*, a global decision is reached as soon as *q* out of the *N* decision makers agree on an answer. Under the assumption of measurement independence among individuals, we introduce a novel numerical method that allows the analysis of the decision accuracy and time of a group of decision makers. We present the numerical method and show a preliminary sets of results that can be obtained using such a method.

I. INTRODUCTION

This work aims to understand how grouping individual decision makers (DM) affects the speed and accuracy with which these individuals reach a collective decision. 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 the family of the *q out of N* aggregations rules. Loosely speaking, the fusion center makes its global decision as soon as one hypothesis gets q votes in its favor. For q ranging from 1 to $\lfloor N/2 \rfloor + 1$, and for distinct sizes of the group of DMs, we study the performance of the decentralized 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.

This work has been supported in part by AFOSR MURI Award F49620-02-1-0325.

S. H. Dandach and R. Carli and F. Bullo are with the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, Santa Barbara, CA 93106, USA, {sandra|carli|rug|bullo}@engr.ucsb.edu.

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 for a wide set of values of q and for varying group sizes. This represents the major contribution of this paper.

We provide in the first part of this work expressions that relate the probabilities of a network making a decision as a function of the network size and fusion rule. In the second part of this work, we use these expressions to characterize the tradeoff between the accuracy and the expected time and the group size and the chosen aggregation rule. For illustration, we consider a discrete distribution of the Koopman-Darmoi-Pitman form. We find that, for a fixed value of N and for increasing values of q , the accuracy improves and the expected number of observations required to provide a decision increases. Similarly, for a fixed value of q and for increasing values of N , the accuracy improves and the expected number of observations required to provide a decision increases. In other words, the greater the size of the group of DMs and the value of q are, the more accurate and slower the decentralized SPRT algorithm is.

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, i.e., the decentralized SPRT with the *q out of N* aggregation rules. 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 $x(1), x(2), x(3), \dots$, 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 t -th observation $x(t)$ as

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

The sum of the log-likelihoods up to k is denoted by $\Lambda(t) = \sum_{j=1}^t \lambda(j)$. For the SPRT, sampling continues as long as $\eta_0 < \Lambda(t) < \eta_1$, where η_0 and η_1 are two pre-assigned thresholds. Sampling is stopped the first time this inequality is violated. If $\Lambda(t) < \eta_0$, then we decide for $\theta = \theta_0$. If $\Lambda(t) > \eta_1$, then we decide 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 obtains $\mathbb{P}[\text{say}H_0|H_1] \leq p_m$ 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).

Observe that the expressions in Equation (2) are exact if, upon stopping with t observations, either $\Lambda(t) = \eta_0$ or $\Lambda(t) = \eta_1$. However some overshooting of the boundaries generally occurs. The distribution of the overshooting $\Lambda(t) - \eta_1$ and $\eta_0 - \Lambda(t)$, quantify the accuracy of the approximations introduced by formulas (2).

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

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 want 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 $t \in \mathbb{N}^1$. Indeed in such case, if $p_{i|j}(t)$ denotes the probability of deciding H_i at time k , given that H_j is the correct hypothesis, i.e., $p_{i|j}(t) := \mathbb{P}[\text{say}H_i|H_j, T = t]$, from standard probabilistic arguments it follows that

$$\mathbb{P}[\text{say}H_1|H_0] = \sum_{t=1}^{\infty} p_{1|0}(t) \quad (3)$$

$$\mathbb{E}[T|H_0] = \sum_{t=1}^{\infty} (p_{0|0}(t) + p_{1|0}(t))t. \quad (4)$$

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 the 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

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

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

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

One can easily see that sampling will continue as long as

$$\frac{\eta_0 + t(A(\theta_1) - A(\theta_0))}{B(\theta_1) - B(\theta_0)} < \sum_{i=1}^t x(i) < \frac{\eta_1 + t(A(\theta_1) - A(\theta_0))}{B(\theta_1) - B(\theta_0)} \quad (5)$$

for $B(\theta_1) - B(\theta_0) > 0$; if $B(\theta_1) - B(\theta_0) < 0$ the inequalities would be reversed. Observe that $\sum_{i=1}^t x(i)$ is an integer number. Now let $\bar{\eta}_0(t)$ be the smallest integer greater than $\{\eta_0 + t(A(\theta_1) - A(\theta_0))\} / (B(\theta_1) - B(\theta_0))$ and let $\bar{\eta}_1(t)$ be the largest integer smaller than $\{\eta_1 + t(A(\theta_1) - A(\theta_0))\} / (B(\theta_1) - B(\theta_0))$. Sampling will continue as long as $\bar{\eta}_0(t) \leq \mathcal{X}(t) \leq \bar{\eta}_1(t)$ where $\mathcal{X}(t) =$

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

$\sum_{i=1}^t x(i)$. Now suppose that, for any $\ell \in [\bar{\eta}_0(t), \bar{\eta}_1(t)]$ the probability $\mathbb{P}[\mathcal{X}(t) = \ell]$ is known. Then we have

$$\mathbb{P}[\mathcal{X}(t+1) = \ell | H_i] = \sum_{j=\bar{\eta}_0(t)}^{\bar{\eta}_1(t)} f(\ell - j; \theta_i) \mathbb{P}[\mathcal{X}(t) = j | H_i], \quad (6)$$

and

$$\begin{aligned} p_{1|i}(t+1) &= \sum_{j=\bar{\eta}_0(t)}^{\bar{\eta}_1(t)} \sum_{r=\bar{\eta}_1(t)-j+1}^{\infty} \mathbb{P}[\mathcal{X}(t) = j | H_i] f(r; \theta_i) \\ p_{0|i}(t+1) &= \sum_{j=\bar{\eta}_0(t)}^{\bar{\eta}_1(t)} \sum_{r=-\infty}^{\bar{\eta}_0(t)-j-1} \mathbb{P}[\mathcal{X}(t) = j | H_i] f(r; \theta_i). \end{aligned} \quad (7)$$

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

III. DECENTRALIZED SPRT: PROBLEM SETUP

The general setting of the hypothesis testing problem we analyze in this paper is described as follows. 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 according to a q out of N counting rule: roughly speaking as soon as one of the two hypothesis, say H_i , receives q local decisions in its favor, the fusion center globally decides in favor of H_i . We refer to the above framework as to the q out of N decentralized SPRT.

Next, we formally describe our setup. Let N denote the size of the group of decision makers and let q be a positive integer such that $1 \leq q \leq \lfloor N/2 \rfloor + 1$. Then the q out of N decentralized SPRT is defined as follows:

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

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

State iteration (Decision Makers) : At time $t \in \mathbb{N}$, for each i , if $\Lambda_i(t-1) \in [\eta_0, \eta_1]$, 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

$$\lambda_i(t) = \log \frac{f(x_i(t), \theta_1)}{f(x_i(t), \theta_0)}. \quad (8)$$

(2) DM i updates the variable Λ_i by

$$\Lambda_i(t) = \Lambda_i(t-1) + \lambda_i(t). \quad (9)$$

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

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

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

State iteration (Fusion Center) : If at time $t \in \mathbb{N}$ the fusion center has not yet provided a global decision, then it performs two actions in the following order:

(1) it updates the variables $Count_+$ and $Count_-$ according to $Count_+(t) = Count_+(t-1) + n_+(t)$ and $Count_-(t) = Count_-(t-1) + n_-(t)$ where the variables $n_+(t)$ and $n_-(t)$ denote, respectively, the number of messages equal to $+1$ and -1 received by the fusion center at time t .

(2) it checks if one of the following two situations is verified

$$(i) \begin{cases} Count_+(t) > Count_-(t) \\ Count_+(t) \geq q \end{cases} \quad (10)$$

$$(ii) \begin{cases} Count_+(t) < Count_-(t) \\ Count_-(t) \geq q \end{cases} \quad (11)$$

If (i) is verified the fusion center globally decides in favor of H_1 , while if (ii) is verified the fusion center globally decides in favor of H_0 . Once the fusion center has provided a global decision the q out of N decentralized SPRT stops.

We introduce now four notational definitions that will be useful throughout all the paper. Given a group of N decision makers and a positive integer q , $1 \leq q \leq \lfloor N/2 \rfloor + 1$, we denote by

- (i) T the random variable accounting for the number of iterations required by the q out of N decentralized SPRT to provide a decision, i.e.,

$$T = \min \{t : \text{either case (10) or case (11) is satisfied}\}$$

- (ii) $p_{i|j}(t; N, q)$ the probability of deciding, at time t , H_i given that H_j is correct, i.e.,

$$p_{i|j}(t; N, q) := \mathbb{P}[\text{Group of } N \text{ DMs says } H_i | H_j, q, T = t]; \quad (12)$$

- (iii) $p_{m;N,q}$ and $p_{f;N,q}$ the probability of mis-detection and of false-alarm, respectively, i.e.,

$$\begin{aligned} p_{m;N,q} &:= \sum_{t=1}^N p_{0|1}(t; N, q) \\ p_{f;N,q} &:= \sum_{t=1}^N p_{1|0}(t; N, q); \end{aligned} \quad (13)$$

- (iv) $\mathbb{E}[T|H_i, N, q]$ the average number of iterations required by the algorithm to provide a decision, given that H_i is the correct hypothesis, i.e.,

$$\mathbb{E}[T|H_i, N, q] := \sum_{t=1}^{\infty} (p_{0|i}(t; N, q) + p_{1|i}(t; N, q))t. \quad (14)$$

Observe that $p_{i|j}(t; 1, 1)$ coincides with the probability $p_{i|j}(t)$ defined in Section II-B when only one DM is running SPRT. For ease of notation we denote by $p_{i|j}(t)$ the case when $N = q = 1$, i.e., $p_{i|j}(t; 1, 1)$. Moreover, when there is no risk of confusion, we might drop some arguments from the definitions in (ii), (iii), (iv) (e.g. denote $p_{i|j}(t; N)$ rather than $p_{i|j}(t; N, q)$). We are now ready to formulate the problem we address in this paper.

Problem III.1 *Given a group of N decision makers running the q out of N decentralized SPRT algorithm to decide between two hypothesis H_0 and H_1 , with two pre-assigned values for the thresholds η_0 and η_1 , estimate the probabilities of mis-detection and of false-alarm of the network, i.e., $p_{m;N,q}$, $p_{f;N,q}$ respectively, and the average number of iterations required to provide a decision, i.e., $\mathbb{E}[T|H_i, N, q]$.*

We will focus on the above problem in the next two Sections, both through theoretical and numerical results. We end this Section by stating a desirable stopping property of the q out of N decentralized SPRT algorithm, when N is odd and $1 \leq q \leq \lfloor N/2 \rfloor + 1$.

Proposition III.2 *Let $N \in \mathbb{N}$ be odd and let $1 \leq q \leq \lfloor N/2 \rfloor + 1$. Then the q out of N decentralized SPRT algorithm provides a decision almost surely.*

Proof: It is well known that the classic SPRT reviewed in Section II-A provides a decision almost surely. Since in the q out of N decentralized SPRT algorithm each decision maker run a SPRT algorithm independently from the other decision makers, there exists almost surely a N -tuple $(t_1, \dots, t_N) \in \mathbb{N}^N$ such that the i -th decision maker provides its decision at time t_i . Let $\bar{t} := \max\{t_i : i \in \{1, \dots, N\}\}$. Since N is odd, then $\text{Count}_+(\bar{t}) \neq \text{Count}_-(\bar{t})$. Moreover since $q \leq \lfloor N/2 \rfloor + 1$ and $\text{Count}_+(\bar{t}) + \text{Count}_-(\bar{t}) = N$,

either $\text{Count}_+(\bar{t}) \geq q$ or $\text{Count}_-(\bar{t}) \geq q$ holds true. Hence one of the two conditions in (11) will hold. ■

In the rest of the paper we will assume the following two properties.

Assumption III.1 *The size of the group of DMs, i.e., N , is odd and the integer q is such that $1 \leq q \leq \lfloor N/2 \rfloor + 1$.*

IV. THE q out of N decentralized SPRT: ANALYSIS OF ACCURACY AND DECISION TIME

The goal of this section is to analyze the accuracy and the expected decision time for the q out of N decentralized SPRT. The key idea is to provide an efficient method to compute the probabilities $p_{i|j}(t; N, q)$, $i, j = 0, 1$. These probabilities, using Equations (13), (14), will also allow to estimate the accuracy and the expected time.

In Section II-B we illustrated a recursive procedure to calculate the probabilities of giving the right and wrong decision, for the standard SPRT algorithm with only one decision maker. Intuitively one would expect that it is possible to compute the probabilities $\{p_{i|j}(t; N, q)\}_{t=1}^{\infty}$, $i, j = 0, 1$ in an analogous way. Indeed, suppose X is a random variable of the Koopman-Darmoi-Pitman form and assume that, for a given $t \in \mathbb{N}$, the probability $\mathbb{P}[\Lambda_1(t) = \ell_1, \dots, \Lambda_N(t) = \ell_N]$ is known for any N -tuple $(\ell_1, \dots, \ell_N) \in [\bar{\eta}_0(t), \bar{\eta}_1(t)]^N$, where $\bar{\eta}_0(t)$ and $\bar{\eta}_1(t)$ are defined as in Section II-B. Then, for any N -tuple $(\bar{\ell}_1, \dots, \bar{\ell}_N) \in [\bar{\eta}_0(t+1), \bar{\eta}_1(t+1)]^N$, one can compute the probability $\mathbb{P}[\Lambda_1(t+1) = \bar{\ell}_1, \dots, \Lambda_N(t+1) = \bar{\ell}_N]$ according to the formula

$$\begin{aligned} \mathbb{P}[\Lambda_1(t+1) = \bar{\ell}_1, \dots, \Lambda_N(t+1) = \bar{\ell}_N] &= \quad (15) \\ &\sum_{j_1=\bar{\eta}_0(t)}^{\bar{\eta}_1(t)} \dots \sum_{j_N=\bar{\eta}_0(t)}^{\bar{\eta}_1(t)} \mathbb{P}[\Lambda_1(t) = j_1, \dots, \Lambda_N(t) = j_N] \times \\ &\times \mathbb{P}[\lambda_1(t+1) = \bar{\ell}_1 - j_1, \dots, \lambda_N(t+1) = \bar{\ell}_N - j_N] \end{aligned}$$

The above formula can be viewed as the N -dimensional extension of the formula in Equation (6). Clearly, providing the corresponding N -dimensional extensions for the formulas in Equation (7), one can also calculate the probabilities $p_{0|j}(t+1; N, q)$ and $p_{1|j}(t+1; N, q)$. In this way, starting from the initial condition $\mathbb{P}[\Lambda_1(0) = 0, \dots, \Lambda_N(0) = 0] = 1$, one can iteratively compute all the elements of the set $\{p_{i|j}(t; N, q)\}_{t=1}^{\infty}$, $i, j = 0, 1$. Observe that, following the above approach, the number of variables that must be kept in memory at any time t is at least $[\bar{\eta}_1^{(t)} - \bar{\eta}_0^{(t)} + 1]^N$; this quantity increases exponentially with the size of the group of DMs and this makes the computational effort required for the computation of the quantities $\{p_{i|j}(t; N, q)\}_{t=1}^{\infty}$, $i, j = 0, 1$ intractable even for small N .

In what follows we provide a more efficient way of computing the probabilities $\{p_{i|j}(t; N, q)\}_{t=1}^{\infty}$, $i, j = 0, 1$, that requires keeping in memory a number of variables which scales only linearly with the number of DMs. 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 assume that they have been computed by the procedures illustrated in Section II-B.

Let us start by introducing some useful quantities.

Define the probability function $\alpha : \mathbb{N} \times \{0, \dots, q-1\} \times \{0, \dots, q-1\} \rightarrow [0, 1]$ as follows: given a group of $s_0 + s_1$ decision makers, $\alpha(t, s_0, s_1)$ is the probability that

- (i) all the $s_0 + s_1$ decision makers have provided a decision up to time t ; and
- (ii) considering the variables $Count_-$ and $Count_+$ restricted to this group of $s_0 + s_1$ decision makers, $Count_-(t) = s_0$ and $Count_+(t) = s_1$.

Define the probability function $\bar{\alpha} : \mathbb{N} \times \{q, \dots, \lfloor N/2 \rfloor\} \rightarrow [0, 1]$ as follows: given a group of $2s$ decision makers, $\bar{\alpha}(t, s)$ is the probability that

- (i) all the $2s$ decision makers have provided a decision up to time t ; and
- (ii) there exists $\bar{\tau} \leq t$ such that, considering the variables $Count_-$ and $Count_+$ restricted to this group of $N - (s_0 + s_1)$ decision makers
 - $Count_-(\bar{\tau} - 1) < q$ and $Count_+(\bar{\tau} - 1) < q$;
 - $Count_-(\tau) = Count_+(\tau) \geq q$ for all $\tau \geq \bar{\tau}$.

Define the probability function $\beta_{1|j} : \mathbb{N} \times \{0, \dots, q-1\} \times \{0, \dots, q-1\} \rightarrow [0, 1]$, $j = 0, 1$ as follows: given a group of $N - (s_0 + s_1)$ decision makers $\beta_{1|j}(t, s_0, s_1)$ is the probability that

- (i) no decision makers have provided a decision up to time $t - 1$; and
- (ii) considering the variables $Count_-$ and $Count_+$ restricted to this group of $N - (s_0 + s_1)$ decision makers, $Count_-(t) + s_0 < Count_+(t) + s_1$, and $Count_+(t) \geq q$.

In a similar way it is possible to define also the probabilities $\beta_{0|j}$, $j = 0, 1$.

Define the probability function $\bar{\beta}_{1|j} : \mathbb{N} \times \{q, \dots, \lfloor N/2 \rfloor\} \rightarrow [0, 1]$, $j = 0, 1$ as follows: given a group of $N - 2s$ decision makers $\bar{\beta}_{1|j}(t, s)$ is the probability that

- (i) no decision makers have provided a decision up to time $t - 1$; and
- (ii) at time t the number of decision makers providing a decision in favor of H_1 is strictly greater of the number of decision makers providing a decision in favor of H_0 .

In a similar way it is possible to define also the probabilities $\bar{\beta}_{0|j}$, $j = 0, 1$.

The following Proposition shows how to compute the probabilities $\{p_{i|j}(t; N, q)\}_{t=1}^{\infty}$, $i, j = 0, 1$, starting from the above definitions.

Proposition IV.1 (q out of N rule: a closed form expression) *Consider the decentralized SPRT with the q out of N rule for a group of N decision makers. Under Assumption III.1 and assuming, without loss of generality, that H_1 is the correct hypothesis, then for $t = 1$ we have that*

$$p_{1|1}(1; N, q) = \beta_{1|1}(1, 0, 0), \quad (16)$$

and

$$p_{0|1}(1; N, q) = \beta_{0|1}(1, 0, 0), \quad (17)$$

while, for $t \geq 2$, we have that

$$\begin{aligned} p_{1|1}(t; N, q) &= \\ &= \sum_{s_0=0}^{q-1} \sum_{s_1=0}^{q-1} \binom{N}{s_1 + s_0} \alpha(t-1, s_0, s_1) \beta_{1|1}(t, s_0, s_1) + \\ &\quad + \sum_{s=q}^{\lfloor N/2 \rfloor} \binom{N}{2s} \bar{\alpha}(t-1, s) \bar{\beta}_{1|1}(t, s). \end{aligned} \quad (18)$$

and

$$\begin{aligned} p_{0|1}(t; N, q) &= \\ &= \sum_{s_0=0}^{q-1} \sum_{s_1=0}^{q-1} \binom{N}{s_1 + s_0} \alpha(t-1, s_0, s_1) \beta_{0|1}(t, s_0, s_1) + \\ &\quad + \sum_{s=q}^{\lfloor N/2 \rfloor} \binom{N}{2s} \bar{\alpha}(t-1, s) \bar{\beta}_{0|1}(t, s). \end{aligned} \quad (19)$$

Proof: The proof that formulas (16) and (17) hold true follows trivially from the definition of the quantities $\beta_{1|1}(1, 0, 0)$ and $\beta_{0|1}(1, 0, 0)$.

As far as formulas (18) and (19) are concerned, we provide the proof only for formula (18), since formula (19) can be proved in a similar way. We start by providing three useful definitions.

First, let E_t denote the event that the decentralized SPRT with the q out of N rule provides its decision at time t in favor of H_1 .

Second, for s_0 and s_1 such that $0 \leq s_0, s_1 \leq q-1$, let $E_{s_0, s_1, t}$ denote the event such that

- (i) there are s_0 DMs that have decided in favor of H_0 up to time $t - 1$;
- (ii) there are s_1 DMs that have decided in favor of H_1 up to time $t - 1$;
- (iii) there exist two positive integer number r_0 and r_1 such that
 - $s_0 + r_0 < s_1 + r_1$ and $s_1 + r_1 \geq q$.
 - at time t , r_0 DMs decides in favor of H_0 while r_1 DMs decides in favor of H_1

Third, for $q \leq s \leq \lfloor N/2 \rfloor$, let $E_{s,t}$ denote the event such that

- (i) $2s$ DMs have provided their decision up to time $t - 1$ balancing their decision, i.e., there exists $\bar{\tau} \leq t - 1$ with the properties that, considering the variables $Count_-$ and $Count_+$ restricted to these $2s$ DMs
 - $Count_-(\tau) < q$, $Count_+(\tau) < q$, for $1 \leq \tau \leq \bar{\tau} - 1$;
 - $Count_-(\tau) = Count_+(\tau)$ for $\bar{\tau} \leq \tau \leq t - 1$;
 - $Count_-(t - 1) = Count_+(t - 1) = s$.
- (ii) at time t the number of DMs providing their decision in favor of H_1 is strictly greater than the number of DMs deciding in favor of H_0 .

Observe that

$$E_t = \left(\bigcup_{0 \leq s_0, s_1 \leq q-1} E_{s_0, s_1, t} \right) \cup \left(\bigcup_{q \leq s \leq \lfloor N/2 \rfloor} E_{s,t} \right).$$

Since the events $E_{s_0, s_1, t}$, $0 \leq s_0, s_1 \leq q-1$, and $E_{s, t}$, $q \leq s \leq \lfloor N/2 \rfloor$ are all disjoint and independent, we can write that

$$\mathbb{P}[E_t] = \sum_{0 \leq s_0, s_1 \leq q-1} \mathbb{P}[E_{s_0, s_1, t}] + \sum_{q \leq s \leq \lfloor N/2 \rfloor} \mathbb{P}[E_{s, t}]. \quad (20)$$

Observe that, according to the definitions of $\alpha(t-1, s_0, s_1)$, $\bar{\alpha}(t-1, s)$, $\beta_{1|1}(t, s_0, s_1)$ and $\bar{\beta}_{1|1}(t, s)$, provided above,

$$\mathbb{P}[E_{s_0, s_1, t}] = \binom{N}{s_1 + s_0} \alpha(t-1, s_0, s_1) \beta_{1|1}(t, s_0, s_1) \quad (21)$$

and that

$$\mathbb{P}[E_{s, t}] = \binom{N}{2s} \bar{\alpha}(t-1, s) \bar{\beta}_{1|1}(t, s). \quad (22)$$

Plugging Equations (21) and (22) into Equation (20) concludes the proof of the Theorem. ■

Formulas, similar to (16), (17), (18) and (19), can be provided for computing also the probabilities $\{p_{i|0}(t; N, q)\}_{t=1}^{\infty}$, $i = 0, 1$.

As far as the probabilities $\alpha(t, s_0, s_1)$, $\bar{\alpha}(t, s)$, $\beta_{i|j}(t, s_0, s_1)$, $\bar{\beta}_{i|j}(t, s)$, $i, j = 0, 1$, are concerned, we provide next expressions for calculating them. First we need one more auxiliary definition. Let the probability function of decision up to time t , $\gamma_{i|j} : \mathbb{N} \rightarrow [0, 1]$, be defined 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.,

$$\gamma_{i|j}(t) = \sum_{s=1}^t p_{i|j}(t) = \sum_{s=1}^{t-1} \gamma_{i|j}(t-1) + p_{i|j}(t). \quad (23)$$

Proposition IV.2 Consider a group of N DMs, running the q out of N decentralized SPRT algorithm under Assumption III.1. Moreover, without loss of generality, assume that H_1 is the correct hypothesis. Then the following statements hold.

First, the probability $\alpha(t, s_0, s_1)$ is given by

$$\alpha(t, s_0, s_1) = \binom{s_0 + s_1}{s_0} \gamma_{0|1}^{s_0}(t) \gamma_{1|1}^{s_1}(t). \quad (24)$$

Second the probability $\bar{\alpha}(t, s)$ satisfies the following recursive relation

$$\begin{aligned} \bar{\alpha}(t, s) &= \\ &= \sum_{s_0=0}^{q-1} \sum_{s_1=0}^{q-1} \binom{2s}{s_0 + s_1} \binom{2s - s_0 - s_1}{s - s_0} \times \\ &\quad \times \alpha(t-1, s_0, s_1) p_{0|1}^{s-s_0}(t) p_{1|1}^{s-s_1}(t) + \\ &\quad + \sum_{h=q}^s \binom{2s}{2h} \binom{2s-2h}{s-h} \bar{\alpha}(t-1, h) p_{0|1}^{s-h}(t) p_{1|1}^{s-h}(t) \end{aligned} \quad (25)$$

Third the probability $\beta_{1|1}(t, s_0, s_1)$ is given by

$$\begin{aligned} \beta_{1|1}(t, s_0, s_1) &= \sum_{j=q-s_1}^{N-\bar{s}} \binom{N-\bar{s}}{j} p_{1|1}^j(t) \times \\ &\left\{ \sum_{i=0}^m \binom{N-\bar{s}-j}{i} p_{0|1}^i(t) (1 - \gamma_{1|1}(t) - \gamma_{0|1}(t))^{N-\bar{s}-i-j} \right\} \end{aligned} \quad (26)$$

where $\bar{s} = s_0 + s_1$ and $m = \min\{j + s_1 - s_0 - 1, N - s_0 + s_1 - j\}$.

Finally the probability $\bar{\beta}_{1|1}(t, s)$ is given by

$$\begin{aligned} \bar{\beta}_{1|1}(t, s) &= \sum_{j=1}^{N-2s} \binom{N-2s}{j} p_{1|1}^j(t) \times \\ &\sum_{i=0}^{\bar{m}} \binom{N-2s-j}{i} p_{0|1}^i(t) (1 - \gamma_{1|1}(t) - \gamma_{0|1}(t))^{N-2s-i-j} \end{aligned} \quad (27)$$

where $\bar{m} = \min\{j-1, N-2s-j\}$.

The expressions for $\beta_{0|1}(t, s_0, s_1)$ and $\bar{\beta}_{0|1}(t, s)$ are obtained by exchanging the roles of $p_{1|1}(t)$ with $p_{0|1}(t)$ in Equations (26) and (27).

The proof of the above Proposition is omitted for the lack of space. We refer the interested reader to the document available at [?].

Now we describe some properties of the above expressions in order to assess the computational complexity required by the formulas introduced in Proposition IV.1 to compute $\{p_{i|j}(t; N, q)\}_{t=1}^{\infty}$, $i, j = 0, 1$. From Equations (24), (25), (26) and (27) we observe that

- $\alpha(t, s_0, s_1)$ is a function of $\gamma_{0|1}(t)$ and $\gamma_{1|1}(t)$;
- $\bar{\alpha}(t, s)$ is a function of $\alpha(t-1, s_0, s_1)$, $0 \leq s_0, s_1 \leq q-1$, $p_{0|1}(t)$, $p_{1|1}(t)$ and $\bar{\alpha}(t-1, h)$, $q \leq h \leq s$;
- $\beta_{i|1}(t, s_0, s_1)$, $\bar{\beta}_{i|1}$, $i = 0, 1$, are functions of $p_{0|1}(t)$, $p_{1|1}(t)$, $\gamma_{0|1}(t)$ and $\gamma_{1|1}(t)$.

Moreover from Equation (23) we have that $\gamma_{i|j}(t)$ is a function of $\gamma_{i|j}(t-1)$ and $p_{i|j}(t)$.

Based on the above observations, we deduce that $p_{0|1}(t; N, q)$ and $p_{1|1}(t; N, q)$ can be seen as the output of a dynamical system having the $(\lfloor N/2 \rfloor - q + 3)$ -th dimensional vector with components the variables $\gamma_{0|1}(t-1)$, $\gamma_{1|1}(t-1)$, $\bar{\alpha}(t-1, s)$, $q \leq h \leq \lfloor N/2 \rfloor$ as states and the two dimensional vector with components $p_{0|1}(t)$, $p_{1|1}(t)$, as inputs. As a consequence, it follows that the iterative method we propose to compute $\{p_{i|j}(t; N, q)\}_{t=1}^{\infty}$, $i, j = 0, 1$, requires keeping in memory a number of variables which grows linearly with the number of decision makers, as opposed to the exponential growth of the method suggested in (15).

V. NUMERICAL RESULTS

The goal of this section is to provide some numerical results of the methods we previously described. For illustration we consider a discrete random variable of the Koopman-Darmoi-Pitman form with a binomial distribution. Specifically, in our numerical examples, we perform the following tasks

- we choose a random variable X with a binomial distribution $f(x; \theta)$;

$$f(x; \theta) = \begin{cases} \binom{n}{x} \theta^x (1-\theta)^{n-x} & \text{if } x \in \{0, 1, \dots, n\} \\ 0 & \text{otherwise} \end{cases}$$

where n is a positive integer number whose value we set to be $n = 5$.

- (ii) we choose two hypothesis $H_0 : \theta = 0.5 - \epsilon$ and $H_1 : \theta = 0.5 + \epsilon$, where $\epsilon = 0.05$, and we assume that $H = H_1$ is the correct hypothesis;
- (iii) we fix $p_m = p_f = 0.1$ and we set the values of the thresholds η_0 and η_1 according to Equations (2), i.e., $\eta_0 = \log(1/9)$ and $\eta_1 = \log 9$;
- (iv) we compute the probabilities $\{p_{i|j}(t)\}_{t=1}^{\infty} = \{p_{i|j}(t, 1)\}_{t=1}^{\infty}$, $i, j \in \{0, 1\}$, through the procedure illustrated in Section II-B;
- (v) we consider group of decision makers of odd size with N ranging from 1 to 35 and for each such group we numerically analyzed the q out of N decentralized SPRT algorithm for $q \in \{1, \dots, \lfloor N/2 \rfloor + 1\}$;
- (vi) for a given group of decision makers of size N and for a given $q \in \{1, \dots, \lfloor N/2 \rfloor + 1\}$ we compute the probabilities $\{p_{i|1}(t; N, q)\}_{t=1}^{\infty}$, $i \in \{0, 1\}$ according to the procedure illustrated in Proposition IV.1 and Proposition IV.2;
- (vii) once we have calculated the probabilities $\{p_{i|j}(t; N, q)\}_{t=1}^{\infty}$, $i, j \in \{0, 1\}$, we compute accuracy and expected time for q out of N decentralized SPRT algorithm exploiting the following formulas

$$\mathbb{P}[\text{say } H_i | H_1] = \sum_{t=1}^{\infty} p_{i|j}(t; N, q)$$

and

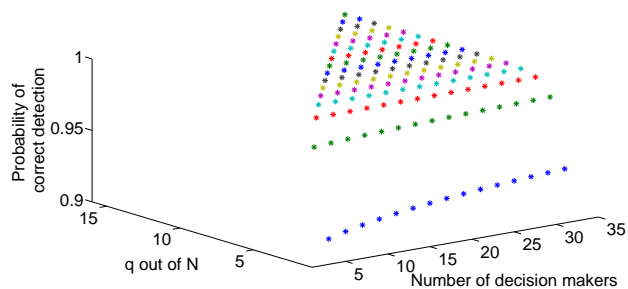
$$\mathbb{E}[T | H_1] = \sum_{t=1}^{\infty} (p_{0|i}(t; N, q) + p_{1|i}(t; N, q))t.$$

The main objective of our numerical examples is to emphasize the tradeoff between accuracy and expected decision time as a function of the number of decision makers and of the q out of N counting rule chosen. The results obtained are reported in Figures 1-5.

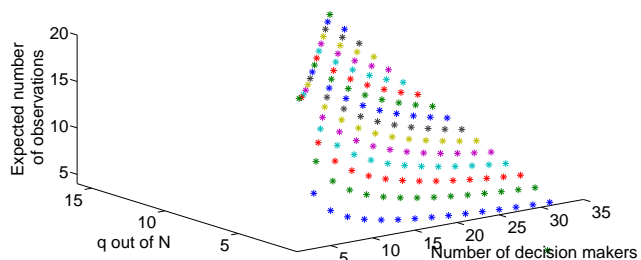
Figures 1 and 2 show the probability of correct detection and the expected decision time as both the group size N and q are varied. For clarity of results, we show in Figures 3- 5, plots for fixed values of N and q .

For a fixed value of N , the accuracy (evaluated in terms of probability of deciding in favor of the correct hypothesis) as well the expected time are increasing with q , i.e., the greater the value of q is, the more accurate and slower the algorithm is (see Figures 3 and 4).

Similarly, for a fixed value of q , the accuracy as well the expected time are increasing with N , i.e., the greater the size of the group of DMS is, the more accurate and slower the algorithm is (see Figures 1, 2 and 5). Specifically in Figure 5 we provide a comparison between two particular values of q , specifically $q = 1$ and $q = \lfloor N/2 \rfloor + 1$. In these specific cases we refer to the q out of N rule as the *fastest* rule if $q = 1$, as the *majority* rule if $q = \lfloor N/2 \rfloor + 1$. 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 time. The fastest rule improves also the accuracy even though this improvement is not as significant



1: This figure shows the probability of correct detection when the q out of N rule is applied on a network of size N .

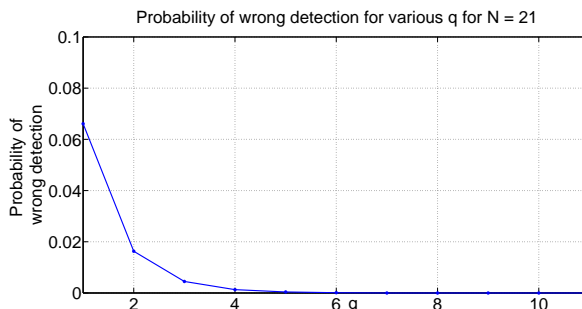


2: This figure shows the expected number of observation to reach a decision when the q out of N rule is applied on a network of size N .

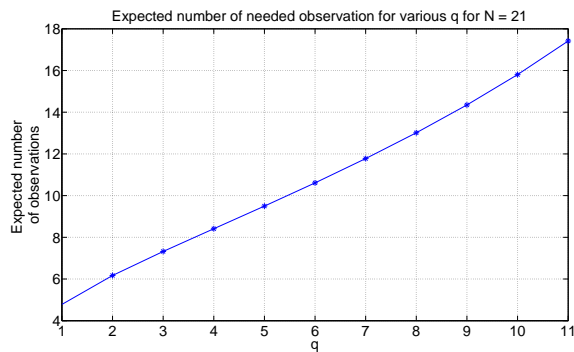
as the one provided by the majority rule. The majority rule, in terms of expected time, exhibits performance which are almost constant as N varies.

VI. CONCLUSION

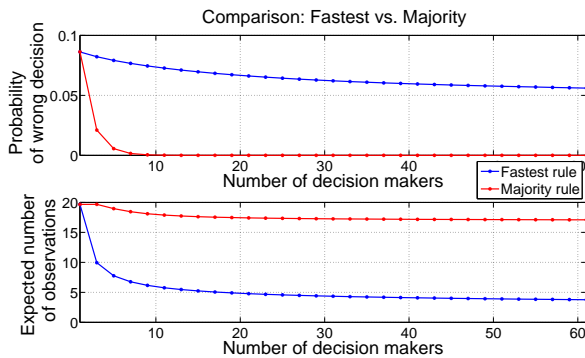
In this work, we presented a complete analysis of how a group of DMs can collectively reach a decision about the correctness of a hypothesis. We presented a numerical method that made it possible to completely analyze and understand interesting fusion rules of the individuals decisions. The general formulas we provided, are useful in analyzing optimality of various aggregation rules, and how the group size and the network desired accuracies dictate



3: This figure shows the probability of correct detection for a networks of fixed sizes $N = 21$, as q in the q out of N rule takes values $q \in \{1, \dots, 11\}$.



4: This figure shows the expected number of observations needed for detection for for a networks of fixed sizes $N = 21$, as q in the q out of N rule takes values $q \in \{1, \dots, 11\}$.



5: This figure shows the probability of correct detection and the expected number of observations needed for detection for different networks of various sizes $N \in \{1, \dots, 31\}$, where q in the q out of N rule takes one of two values $q \in \{1, \lfloor N/2 \rfloor + 1\}$. The plots shows the behavior as N increases.

the better aggregation rule to be used. In fact, in an earlier work, we studied in details how two specific aggregation rules, the *fastest* and the *majority* rule, that both belong to the q out of N aggregation rule. We showed how the better aggregation rule varies with the network size and desired network accuracy. Finally interesting extensions of this work, include allowing communication between the various decision makers and allowing different accuracies between different members of the group.

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.