Accuracy and Decision Time for Sequential Decision Aggregation

Sandra H. Dandach   Ruggero Carli   Francesco Bullo

Abstract

This paper studies prototypical strategies to sequentially aggregate independent decisions. We consider a collection of agents, each performing binary hypothesis testing and each obtaining a decision over time. We assume the agents are identical and receive independent information. Individual decisions are sequentially aggregated via a threshold-based rule. In other words, a collective decision is taken as soon as a specified number of agents report a concordant decision (simultaneous discordant decisions and no-decision outcomes are also handled).

We obtain the following results. First, we characterize the probabilities of correct and wrong decisions as a function of time, group size and decision threshold. The computational requirements of our approach are linear in the group size. Second, we consider the so-called fastest and majority rules, corresponding to specific decision thresholds. For these rules, we provide a comprehensive scalability analysis of both accuracy and decision time. In the limit of large group sizes, we show that the decision time for the fastest rule converges to the earliest possible individual time, and that the decision accuracy for the majority rule shows an exponential improvement over the individual accuracy. Additionally, via a theoretical and numerical analysis, we characterize various speed/accuracy tradeoffs. Finally, we relate our results to some recent observations reported in the cognitive information processing literature.

I. Introduction

A. Problem setup

Interest in group decision making spans a wide variety of domains. Be it in electoral votes in politics, detection in robotic and sensor networks, or cognitive data processing in the human brain, establishing the best strategy or understanding the motivation behind an observed strategy, has been of interest for many researchers. This work aims to understand how grouping individual sequential decision makers 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 and Bayesian learning in social networks.

In our problem, a group of individuals independently decide between two alternative hypothesis, and each individual sends its local decision to a fusion center. The fusion center decides for the whole group as soon as one hypothesis gets a number of votes that crosses a pre-determined threshold. We are interested in relating the accuracy and decision time of the whole population, to the accuracy and decision time of a single individual. We assume that all individuals are independent and identical. That is, we assume that they gather information corrupted by i.i.d. noise and that the same statistical test is used by each individual in the population. The setup of similar problems studied in the literature usually assumes that all individual decisions need to be available to the fusion center, before the latter can reach a final decision. The work presented here relaxes this assumption and the fusion center might provide the global decision much earlier than the all individuals in the group. Researchers in behavioral studies refer to decision making schemes where everyone is given an equal amount of time to respond as the “free response paradigm.” Since the speed of the group’s decision is one of our main concerns, we adjust the analysis in a way that makes it possible to compute the joint probabilities of each decision at each time instant. Such a paradigm is referred to as the “interrogation paradigm.”

B. Literature review

The framework we analyze in this paper is related to the one considered in many papers in the literature, see for instance [1], [2], [3], [4], [5], [6], [7], [8], [9] and references therein. The focus of these works is mainly two-fold. First, researchers in the fields aim to determine which type of information the decision makers should send to the

This work has been supported in part by AFOSR MURI FA9550-07-1-0528.

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|carlirug|bullo}@engr.ucsb.edu.

fusion center. Second, many of the studies concentrate on computing optimal decision rules both for the individual decision makers and the fusion center where optimality refers to maximizing accuracy. One key implicit assumption made in numerous works, is that the aggregation rule is applied by the fusion center only after all the decision makers have provided their local decisions.

Tsitsiklis in [1] studied the Bayesian decision problem with a fusion center and showed that for large groups identical local decision rules are asymptotically optimal. Varshney in [2] proved that when the fusion rules at the individuals level are non-identical, threshold rules are the optimal rules at the individual level. Additionally, Varshney proved that setting optimal thresholds for a class of fusion rules, where a decision is made as soon as a certain number q out of the N group members decide, requires solving a number of equations that grows exponentially with the group size. The fusion rules that we study in this work fall under the q out of N class of decision rules. Finally, Varshney proved that this class of decision rules is optimal for identical local decisions.

C. Contributions

The contributions of this paper are three-folds. First, we introduce a recursive approach to characterize the probabilities of correct and wrong decisions for a group of sequential decision makers (SDMs). These probabilities are computed as a function of time, group size and decision threshold. The key idea is to relate the decision probability for a group of size N at each time t, to the decision probability of an individual SDM up to that time t, in a recursive manner. Our proposed method has many advantages. First, our method has a numerical complexity that grows only linearly with the number of decision makers. Second, our method is independent of the specific decision making test adopted by the SDMs and requires knowledge of only the decision probabilities of the SDMs as a function of time. Third, our method allows for asynchronous decision times among SDMs. To the best of our knowledge, the performance of sequential aggregation schemes for asynchronous decisions has not been previously studied.

Second, we consider the so-called *fastest* and *majority* rules corresponding, respectively, to the decision thresholds $q = 1$ and $q = \lceil N/2 \rceil$. For these rules we provide a comprehensive scalability analysis of both accuracy and decision time. Specifically, in the limit of large group sizes, we provide exact expressions for the expected decision time and the probability of wrong decision for both rules, as a function of the decision probabilities of each SDM. For the *fastest* rule we show that the group decision time converges to the earliest possible decision time of an individual SDM, i.e., the earliest time for which the individual SDM has a non-zero decision probability. Additionally, the *fastest* rule asymptotically obtains the correct answer almost surely, provided the individual SDM is more likely to make the correct decision, rather than the wrong decision, at the earliest possible decision time. For the *majority* rule we show that the probability of wrong decision converges exponentially to zero if the individual SDM has a sufficiently small probability of wrong decision. Additionally, the decision time for the *majority* rule is related to the earliest time at which the individual SDM is more likely to give a decision than to not give a decision. This scalability analysis relies upon novel asymptotic and monotonicity results of certain binomial expansions.

As third main contribution, using our recursive method, we present a comprehensive numerical analysis of sequential decision aggregation based on the q *out of* N rules. As model for the individual SDMs, we adopt the sequential probability ratio test (SPRT), which we characterize as an absorbing Markov chain. First, for the *fastest* and *majority* rules, we report how accuracy and decision time vary as a function of the group size and of the SPRT decision probabilities. Second, in the most general setup, we report how accuracy and decision time vary monotonically as a function of group size and decision threshold. Additionally, we compare the performance of fastest versus majority rules, at fixed group accuracy. We show that the best choice between the fastest rule and the majority rule is a function of group size and group accuracy. Our numerical results illustrate why the design of optimal aggregation rules is a complex task [10]. Finally, we discuss possible relationships between our analysis of sequential decision aggregation and mental behavior documented in the cognitive psychology and neuroscience literature [11], [12], [13], [14].

Finally, we draw some qualitative lessons about sequential decision aggregation from our mathematical analysis. Surprisingly, our results show that the accuracy of a group is not necessarily improved over the accuracy of an individual. In aggregation based on the *majority* rule, it is true that group accuracy is (exponentially) better than individual accuracy; decision time, however, converges to a constant value for large group sizes. Instead, if a quick decision time is desired, then the *fastest* rule leads, for large group sizes, to decisions being made at the earliest possible time. However, the accuracy of fastest aggregation is not determined by the individual accuracy (i.e., the time integral of the probability of correct decision over time), but is rather determined by the individual accuracy at a specific time instant, i.e., the probability of correct decision at the earliest decision time. Accuracy at this special time might be arbitrarily bad especially for ”asymmetric” decision makers (e.g., SPRT with asymmetric thresholds). Arguably, these detailed results for *fastest* and *majority* rules, $q = 1$ and $q = \lfloor N/2 \rfloor$ respectively, are indicative of the accuracy and decision time performance of aggregation rules for small and large thresholds, respectively.

D. Decision making in cognitive psychology

An additional motivation to study sequential decision aggregation is our interest in sensory information processing systems in the brain. There is a growing belief among neuroscientists [12], [13], [14] that the brain normally engages in an ongoing synthesis of streams of information (stimuli) from multiple sensory modalities. Example modalities include vision, auditory, gustatory, olfactory and somatosensory. While many areas of the brain (e.g., the primary projection pathways) process information from a single sensory modality, many nuclei (e.g., in the Superior Colliculus) are known to receive and integrate stimuli from multiple sensory modalities. Even in these multi-modal sites, a specific stimulus might be dominant. Multi-modal integration is indeed relevant when the response elicited by stimuli from different sensory modalities is statistically different from the response elicited by the most effective of those stimuli presented individually. (Here, the response is quantified in the number of impulses from neurons.) Moreover, regarding data processing in these multi-modal sites, the procedure with which stimuli are processed changes depending upon the intensity of each modality-specific stimulus.

In [12], Werner et al. study a human decision making problem with multiple sensory modalities. They present examples where accuracy and decision time depend upon the strength of the audio and visual components in audio-visual stimuli. They find that, for intact stimuli (i.e., noiseless signals), the decision time improves in multi-modal integration (that is, when both stimuli are simultaneously presented) as compared with uni-sensory integration. Instead, when both stimuli are degraded with noise, multi-modal integration leads to an improvement in both accuracy and decision time. Interestingly, they also identify circumstances for which multi-modal integration leads to performance degradation: performance with an intact stimulus together with a degraded stimulus is sometimes worse than performance with only the intact stimulus.

Another point of debate among cognitive neuroscientists is how to characterize uni-sensory versus multi-modal integration sites. Neuro-physiological studies have traditionally classified as multi-modal sites where stimuli are enhanced, that is, the response to combined stimuli is larger than the sum of the responses to individual stimuli. Recent observations of suppressive responses in multi-modal sites has put this theory to doubt; see [13], [14] and references therein. More specifically, studies have shown that by manipulating the presence and informativeness of stimuli, one can affect the performance (accuracy and decision time) of the subjects in interesting, yet not well understood ways. We envision that a more thorough theoretical understanding of sequential decision aggregation will help bridge the gap between these seemingly contradicting characterization of multi-modal integration sites.

As a final remark about uni-sensory integration sites, it is well known [15] that the cortex in the brain integrates information in *neural groups* by implementing a *drift-diffusion model*. This model is the continuous-time version of the so-called sequential probability ratio test (SPRT) for binary hypothesis testing. We will adopt the SPRT model for our numerical results.

E. Organization

We start in Section II by introducing the problem setup. In Section III we present the numerical method that allows us to analyze the decentralized Sequential Decision Aggregation (SDA) problem; We analyze the two proposed rules in Section IV. We also present the numerical results in Section V. Our conclusions are stated in Section VI. The appendices contain some results on binomial expansions and on the SPRT.

II. Models of sequential aggregation and problem statement

In this section we introduce the model of sequential aggregation and the analysis problem we want to address. Specifically in Subsection II-A we review the classical sequential binary hypothesis testing problem and the notion of *sequential decision maker*, in Subsection II-B we define the *q out of N sequential decisions aggregation* setting and, finally, in Subsection II-C, we state the problem we aim to solve.

A. Sequential decision maker

The classical binary sequential decision problem is posed as follows.

Let H denote a hypothesis which takes on values H_0 and H_1. Assume we are given an individual (called *sequential decision maker (SDM)* hereafter) who repeatedly observes at time $t = 1, 2, \dots,$ a random variable X taking values in some set $\mathcal{X}$ with the purpose of deciding between H_0 and H_1. Specifically the SDM takes the observations $x(1), x(2), x(3), \dots$, until it provides its final decision at time τ, which is assumed to be a stopping time for the sigma field sequence generated by the observations, and makes a final decision δ based on the observations up to time τ. The stopping rule together with the final decision rule represent the decision policy of the SDM. The standing assumption is that the conditional joint distributions of the individual observations under each hypothesis are known to the SDM.

In our treatment, we do not specify the type of decision policy adopted by the SDM. A natural way to keep our presentation as general as possible, is to refer to a probabilistic framework that conveniently describes the sequential decision process generated by any decision policy. Specifically, given the decision policy γ, let $\chi_0^{(\gamma)}$ and $\chi_1^{(\gamma)}$ be two random variables defined on the sample space $\mathbb{N} \times \{0, 1\} \cup \{?\}$ such that, for $i, j \in \{0, 1\}$,

- $\{\chi_j^{(\gamma)} = (t, i)\}$ represents the event that the individual decides in favor of H_i at time t given that the true hypothesis is H_j; and
- $\{\chi_j^{(\gamma)} =?\}$ represents the event that the individual never reaches a decision given that H_j is the correct hypothesis.

Accordingly, define $p_{i|j}^{(\gamma)}(t)$ and $p_{\text{nd}|j}^{(\gamma)}$ to be the probabilities that, respectively, the events $\{\chi_j^{(\gamma)} = (t, i)\}$ and $\{\chi_0^{(\gamma)} =?\}$ occur, i.e,

$$p_{i|j}^{(\gamma)}(t) = \mathbb{P}[\chi_j^{(\gamma)} = (t, i)] \qquad \text{and} \qquad p_{\text{nd}|j}^{(\gamma)} = \mathbb{P}[\chi_j^{(\gamma)} =?].$$

Then the sequential decision process induced by the decision policy γ is completely characterized by the following two sets of probabilities

$$\left\{p_{\text{nd}|0}^{(\gamma)}\right\} \cup \left\{p_{0|0}^{(\gamma)}(t), p_{1|0}^{(\gamma)}(t)\right\}_{t\in\mathbb{N}} \qquad \text{and} \qquad \left\{p_{\text{nd}|1}^{(\gamma)}\right\} \cup \left\{p_{0|1}^{(\gamma)}(t), p_{1|1}^{(\gamma)}(t)\right\}_{t\in\mathbb{N}}, \tag{1}$$

where, clearly $p_{\text{nd}|0}^{(\gamma)} + \sum_{t=1}^{\infty} \left(p_{0|0}^{(\gamma)}(t) + p_{1|0}^{(\gamma)}(t)\right) = 1$ and $p_{\text{nd}|1}^{(\gamma)} + \sum_{t=1}^{\infty} \left(p_{0|1}^{(\gamma)}(t) + p_{1|1}^{(\gamma)}(t)\right) = 1$. In what follows, while referring to a SDM running a sequential distributed hypothesis test with a pre-assigned decision policy, we will assume that the above two probabilities sets are known. From now on, for simplicity, we will drop the superscript (γ).

Together with the probability of no-decision, for $j \in \{0, 1\}$ we introduce also the probability of correct decision $p_{\text{c}|j} := \mathbb{P}[\text{say } H_j \mid H_j]$ and the probability of wrong decision $p_{\text{w}|j} := \mathbb{P}[\text{say } H_i,\ i \neq j \mid H_j]$, that is,

$$p_{\text{c}|j} = \sum_{t=1}^{\infty} p_{j|j}(t) \qquad \text{and} \qquad p_{\text{w}|j} = \sum_{t=1}^{\infty} p_{i|j}(t),\ \ i \neq j.$$

It is worth remarking that in most of the binary sequential decision making literature, $p_{\text{w}|1}$ and $p_{\text{w}|0}$ are referred as, respectively, the *mis-detection* and *false-alarm* probabilities of error.

Below, we provide a formal definition of two properties that the SDM might or might not satisfy.

Definition II.1 *For a SDM with decision probabilities as in* (1)*, the following properties may be defined:*

(i) *the SDM has* almost-sure decisions *if, for* $j \in \{0, 1\}$,

$$\sum_{t=1}^{\infty} \left(p_{0|j}(t) + p_{1|j}(t)\right) = 1, \quad \textit{and}$$

(ii) *the SDM has* finite expected decision time *if, for* $j \in \{0, 1\}$,

$$\sum_{t=1}^{\infty} t \left(p_{0|j}(t) + p_{1|j}(t)\right) < \infty.$$

One can show that the finite expected decision time implies almost-sure decisions.

We conclude this section by briefly discussing examples of sequential decision makers. The classic model is the SPRT model, which we discuss in some detail in the example below and in Section V. Our analysis, however, allows for arbitrary sequential binary hypothesis tests, such as the SPRT with time-varying thresholds [16], constant false alarm rate tests [17], and generalized likelihood ratio tests. Response profiles arise also in neurophysiology, e.g., [18] presents neuron models with a response that varies from unimodal to bimodal depending on the strength of the received stimulus.

Example II.2 (Sequential probability ratio test (SPRT)) In the case the observations taken are independent, conditioned on each hypothesis, a well-known solution to the above binary decision problem is the so-called *sequential probability ratio test (SPRT)* that we review in Section V. A SDM implementing the SPRT test has both the *almost-sure decisions* and *finite expected decision time* properties. Moreover the SPRT test satisfies the following optimality property: among all the sequential tests having pre-assigned values of *mis-detection* and *false-alarm* probabilities of error, the SPRT is the test that requires the smallest expected number of iterations for providing a solution.

In Appendices B.1 and B.2 we review the methods proposed for computing the probabilities $\{p_{i|j}(t)\}_{t\in\mathbb{N}}$ when the SPRT test is applied, both in the case X is a discrete random variable and in the case X is a continuous random variable. For illustration purposes, we provide in Figure 1 the probabilities $p_{i|j}(t)$ when $j = 1$ for the case when X is a continuous random variable with a continuous distribution (Gaussian). We also note that $p_{i|j}(t)$ might have various interesting distributions.

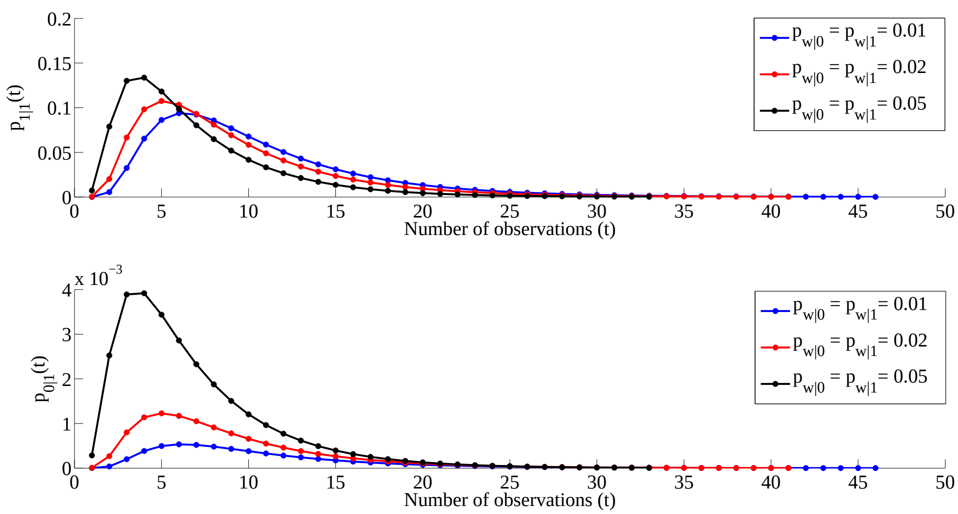

Fig. 1. This figure illustrates a typical unimodal set of decision probabilities $\{p_{1|1}(t)\}_{t\in\mathbb{N}}$ and $\{p_{0|1}(t)\}_{t\in\mathbb{N}}$. Here the SDM is implementing the sequential probability ratio test with three different accuracy levels (see Section V for more details).

B. The q out of N decentralized hypothesis testing

The basic framework for the binary hypothesis testing problem we analyze in this paper is the one in which there are N SDMs and one fusion center. The binary hypothesis is denoted by H and it is assumed to take on values H_0 and H_1. Each SDM is assumed to perform individually a binary sequential test; specifically, for $i \in \{1, \dots, N\}$, at time $t \in \mathbb{N}$, SDM i takes the observation $x_i(t)$ on a random variable X_i, defined on some set $\mathcal{X}_i$, and it keeps observing X_i until it provides its decision according to some decision policy γ_i. We assume that

(i) the random variables $\{X_i\}_{i=1}^N$ are identical and independent;
(ii) the SDMs adopt the same decision policy γ, that is, $\gamma_i \cong \gamma$ for all $i \in \{1, \dots, N\}$;

(iii) the observations taken, conditioned on either hypothesis, are independent from one SDM to another;
(iv) the conditional joint distributions of the individual observations under each hypothesis are known to the SDMS.

In particular assumptions (i) and (ii) imply that the N decision processes induced by the N SDMs are all described by the same two sets of probabilities

$$\left\{p_{\mathrm{nd}|0}\right\} \cup \left\{p_{0|0}(t), p_{1|0}(t)\right\}_{t\in\mathbb{N}} \qquad \text{and} \qquad \left\{p_{\mathrm{nd}|1}\right\} \cup \left\{p_{0|1}(t), p_{1|1}(t)\right\}_{t\in\mathbb{N}}. \tag{2}$$

We refer to the above property as *homogeneity* among the SDMs.

Once a SDM arrives to a final local decision, it communicates it to the 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 the hypothesis H_i receives q local decisions in its favor, the fusion center globally decides in favor of H_i. In what follows we refer to the above framework as q *out of* N *sequential decision aggregation* with homogeneous SDMs (denoted as q *out of* N *SDA*, for simplicity).

We describe our setup in more formal terms. Let N denote the size of the group of SDMs and let q be a positive integer such that $1 \leq q \leq N$, then the q *out of* N *SDA* with homogeneous SDMs is defined as follows:

SDMs iteration : For each $i \in \{1, \dots, N\}$, the i-th SDM keeps observing X_i, taking the observations $x_i(1), x_i(2), \dots,$ until time τ_i where it provides its local decision $d_i \in \{0, 1\}$; specifically $d_i = 0$ if it decides in favor of H_0 and $d_i = 1$ if it decides in favor of H_1. The decision d_i is instantaneously communicated (i.e., at time τ_i) to the fusion center.

Fusion center state : The fusion center stores in memory the variables $Count_0$ and $Count_1$, which are initialized to 0, i.e., $Count_0(0) = Count_1(0) = 0$. 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 $Count_0$ and $Count_1$, according to $Count_0(t) = Count_0(t-1) + n_0(t)$ and $Count_1(t) = Count_1(t-1) + n_1(t)$ where $n_0(t)$ and $n_1(t)$ denote, respectively, the number of decisions equal to 0 and 1 received by the fusion center at time t.
(2) it checks if one of the following two situations is verified

$$(i) \begin{cases} Count_1(t) > Count_0(t), \\ Count_1(t) \geq q, \end{cases} \qquad (ii) \begin{cases} Count_1(t) < Count_0(t). \\ Count_0(t) \geq q. \end{cases} \tag{3}$$

If (i) is verified the fusion center globally decides in favor 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 *SDA* algorithm stops.

Remark II.3 (Notes about SDA) (i) Each SDM has in general a non-zero probability of not giving a decision. In this case, the SDM might keep sampling infinitely without providing any decision to the fusion center.
(ii) The fusion center does not need to wait until all the SDM have provided a decision before a decision is reach on the group level, as one of the two conditions (i) or (ii) in equation 3 might be satisfied much before the N SDM provide their decisions.
(iii) While we study in this manuscript the case when a fusion center receives the information from all SDM, we note that a distributed implementation of the SDA algorithm is possible. Analysis similar to the one presented here is possible in that case. □

C. Problem formulation

We introduce now some definitions that will be useful throughout this paper. Given a group of N SDMs running the q *out of* N *SDA* algorithm, $1 \leq q \leq N$, we denote
(i) by T the random variable accounting for the number of iterations required to provide a decision

$$T = \min\{t \mid \text{either } \textit{case} \text{ (i) or } \textit{case} \text{ (ii) in equation (3) is satisfied}\};$$

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

$$p_{i|j}(t; N, q) := \mathbb{P}\left[\text{Group of } N \text{ SDMs says } H_i \mid H_j, q, T = t\right]; \tag{4}$$

(iii) by $p_{\mathrm{c}|j}(N,q)$ and $p_{\mathrm{w}|j}(N,q)$ the probability of correct decision and of wrong decision, respectively, given that H_j is the correct hypothesis, i.e.,

$$p_{\mathrm{c}|j}(N,q) = \sum_{t=1}^{\infty} p_{j|j}(t;N,q) \qquad \text{and} \qquad p_{\mathrm{w}|j}(N,q) = \sum_{t=1}^{\infty} p_{i|j}(t;N,q),\ i \neq j; \tag{5}$$

(iv) by $p_{\mathrm{nd}|j}(N,q)$, $j \in \{0,1\}$, the probability of no-decision given that H_j is the correct hypothesis, i.e.,

$$p_{\mathrm{nd}|j}(N,q) := 1 - \sum_{t=1}^{\infty} \big(p_{0|j}(t;N,q) + p_{1|j}(t;N,q)\big) = 1 - p_{\mathrm{w}|j}(N,q) - p_{\mathrm{c}|j}(N,q); \tag{6}$$

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

$$\mathbb{E}\left[T|H_j,N,q\right] := \begin{cases} \sum_{t=1}^{\infty} t(p_{0|j}(t;N,q) + p_{1|j}(t;N,q)), & \text{if } p_{\mathrm{nd}|j}(N,q) = 0, \\ +\infty, & \text{if } p_{\mathrm{nd}|j}(N,q) > 0. \end{cases} \tag{7}$$

Observe that $p_{i|j}(t;1,1)$ coincides with the probability $p_{i|j}(t)$ introduced in (1). For ease of notation we will continue using $p_{i|j}(t)$ instead of $p_{i|j}(t;1,1)$.

We are now ready to formulate the problem we aim to solve in this paper.

Problem II.4 (Sequential decision aggregation) *Consider a group of N homogeneous SDMs with decision probabilities $\{p_{\mathrm{nd}|0}\} \cup \{p_{0|0}(t), p_{1|0}(t)\}_{t\in\mathbb{N}}$ and $\{p_{\mathrm{nd}|1}\} \cup \{p_{0|1}(t), p_{1|1}(t)\}_{t\in\mathbb{N}}$. Assume the N SDMs run the q* out of N SDA *algorithm with the purpose of deciding between the hypothesis H_0 and H_1. For $j \in \{0,1\}$, compute the distributions $\{p_{i|j}(t;N,q)\}_{t\in\mathbb{N}}$ as well as the probabilities of correct and wrong decision, i.e., $p_{\mathrm{c}|j}(N,q)$ and $p_{\mathrm{w}|j}(N,q)$, the probability of no-decision $p_{\mathrm{nd}|j}(N,q)$ and the average number of iterations required to provide a decision, i.e., $\mathbb{E}\left[T|H_j,N,q\right]$.*

We will focus on the above problem in the next two Sections, both through theoretical and numerical results. Moreover, in Section IV, we will concentrate on two particular values of q, specifically for $q=1$ and $q=\lfloor N/2 \rfloor + 1$, characterizing the tradeoff between the expected decision time, the probabilities of correct and wrong decision and the size of the group of SDMs. When $q=1$ and $q=\lceil N/2 \rceil$, we will refer to the q *out of* N *rule* as the *fastest rule* and the *majority rule*, respectively. In this case we will use the following notations

$$p^{(\mathrm{f})}_{\mathrm{c}|j}(N) := p_{\mathrm{c}|j}(N;q=1), \qquad p^{(\mathrm{f})}_{\mathrm{w}|j}(N) := p_{\mathrm{w}|j}(N;q=1)$$

and

$$p^{(\mathrm{m})}_{\mathrm{c|j}}(N) := p_{\mathrm{c}|j}(N;q=\lfloor N/2 \rfloor + 1), \qquad p^{(\mathrm{m})}_{\mathrm{w}|j}(N) := p_{\mathrm{w}|j}(N;q=\lfloor N/2 \rfloor + 1).$$

We end this Section by stating two propositions characterizing the *almost-surely decisions* and *finite expected decision time* properties for the group of SDMs.

Proposition II.5 *Consider a group of N SDMs running the q* out of N SDA *algorithm. Let the decision-probabilities of each SDM be as in* (2)*. For $j \in \{0,1\}$, assume there exists at least one time instant $t_j \in \mathbb{N}$ such that both probabilities $p_{0|j}(t_j)$ and $p_{1|j}(t_j)$ are different from zero. Then the group of SDMs has the* almost-sure decision *property if and only if*

(i) *the single SDM has the* almost-sure decision *property;*
(ii) *N is odd; and*
(iii) *q is such that $1 \leq q \leq \lceil N/2 \rceil$.*

Proof: First we prove that if the group of SDMs has the *almost-sure decision* property, then properties (i), (ii) and (iii) are satisfied. To do so, we show that if one between the properties (i), (ii) and (iii) fails then there exists an event of probability non-zero that leads the group to not provide a decision. First assume that the single SDM does not have the *almost-sure decision* property, i.e., $p_{\mathrm{nd}|j} > 0$, $j \in \{0,1\}$. Clearly this implies that the event *"all the SDMs of the group do not provide a decision"* has probability of occurring equal to $p^N_{\mathrm{nd}|j}$ which is strictly greater than zero. Second assume that N is even and consider the event *"at time t_j, $N/2$ SDMs decide in favor of H_0 and $N/2$ SDMs decide in favor of H_1"*. Simple combinatoric and probabilistic arguments show that the probability of this event is $\binom{N}{N/2} p_{0|j}^{N/2} p_{1|j}^{N/2}$, which is strictly greater than zero because of the assumption $p_{0|j}(t_j) \neq 0$ and $p_{1|j}(t_j) \neq 0$. Third assume that $q > \lfloor N/2 \rfloor + 1$. In this case we consider the event *"at time t_j, $\lceil N/2 \rceil$ SDMs decide in favor of H_0 and $\lfloor N/2 \rfloor$ SDMs decide in favor of H_1"* that, clearly, leads the group of SDMs to not provide a global decision for any $q > \lfloor N/2 \rfloor + 1$. Similarly to the previous case, we have that the probability of this event is $\binom{N}{\lceil N/2 \rceil} p_{0|j}^{\lceil N/2 \rceil} p_{1|j}^{\lfloor N/2 \rfloor} > 0$.

We prove now that if properties (i), (ii) and (iii) are satisfied then the group of SDMs has the *almost-sure decision* property. Observe that, since each SDM has the *almost-sure decision* property, there exists almost surely a N-tuple $(t_1, \dots, t_N) \in \mathbb{N}^N$ such that the i-th SDM provides its decision at time t_i. Let $\bar{t} := \max\{t_i \mid i \in \{1, \dots, N\}\}$. Since N is odd, then $Count_1(\bar{t}) \neq Count_0(\bar{t})$. Moreover since $q \leq \lfloor N/2 \rfloor + 1$ and $Count_1(\bar{t}) + Count_0(\bar{t}) = N$, either $Count_1(\bar{t}) \geq q$ or $Count_0(\bar{t}) \geq q$ holds true. Hence the fusion center will provide a global decision not later than time $\bar{t}$. ■

Proposition II.6 *Consider a group of N SDMs running the q* out of *N* SDA *algorithm. Let the decision-probabilities of each SDM be as in* (2)*. For $j \in \{0, 1\}$, assume there exists at least one time instant $t_j \in \mathbb{N}$ such that both probabilities $p_{1|j}(t_j)$ and $p_{1|j}(t_j)$ are different from zero. Then the group of SDMs has the* finite expected decision time *property if and only if*

(i) *the single SDM has the* finite expected decision time *property;*
(ii) *N is odd; and*
(iii) *q is such that $1 \leq q \leq \lceil N/2 \rceil$.*

Proof: The proof follows the lines of the proof of the previous proposition. ■

Remark II.7 The existence, for $j \in \{0, 1\}$, of a time t_j such that $p_{0|j}(t_j) \neq 0$ and $p_{1|j}(t_j) \neq 0$, is necessary only for proving the "if" side of the previous propositions. In other words the validity of properties (i), (ii) and (iii) in Proposition II.5 (resp. in Prop. II.6) guarantees that the group of SDMs possesses the *almost-sure decision* property (resp. the *finite expected decision time* property.) □

III. Recursive analysis of the q-out-of-N sequential aggregation rule

The goal of this section is to provide an efficient method to compute the probabilities $p_{i|j}(t; N, q)$, $i, j \in \{0, 1\}$. These probabilities, using equations (5), (6) and (7) will allow us to estimate the probabilities of correct decision, wrong decision and no-decision, as well as the expected number of iterations required to provide the final decision.

We first consider in subsection III-A the case where $1 \leq q \leq \lfloor N/2 \rfloor$; in subsection III-B we consider the case where $\lfloor N/2 \rfloor + 1 \leq q \leq N$.

A. Case $1 \leq q \leq \lfloor N/2 \rfloor$

To present our analysis method, we begin with an informal description of the decision events characterizing the *q out of N* SDA algorithm. Assume that the fusion center provides its decision at time t. This fact implies that neither case (i) nor case (ii) in equation (3) has happened at any time before t. Moreover, two distinct set of events may precede time t, depending upon whether the values of the counters $Count_0$ and $Count_1$ at time $t-1$ are smaller than q or not. In a first possible set of events, say the "simple situation," the counters satisfy $0 \leq Count_0(t-1), Count_1(t-1) \leq q-1$ and, hence, the time t is the first time that at least one of the two counters crosses the threshold q. In a second possible set of events, say the "canceling situation," the counters $Count_0(t-1)$ and $Count_1(t-1)$ are greater than q and, therefore, equal. In the canceling situation, there must exist a time instant $\bar{\tau} \leq t-1$ such that $Count_0(\bar{\tau}-1) < q$, $Count_1(\bar{\tau}-1) < q$ and $Count_0(\tau) = Count_1(\tau) \geq q$ for all $\tau \in \{\bar{\tau}+1, \dots, t-1\}$. In other words, both counters cross the threshold q at the same time instant $\bar{\tau}$ reaching the same value, that is, $Count_0(\bar{\tau}) = Count_1(\bar{\tau})$, and, for time $\tau \in \{\bar{\tau}+1, \dots, t-1\}$, the number $n_0(\tau)$ of SDMs deciding in favor of H_0 at time τ and the number $n_1(\tau)$ of SDMs deciding in favor of H_1 at time τ cancel each other out, that is, $n_0(\tau) = n_1(\tau)$.

In what follows we study the probability of the simple and canceling situations. To keep track of both possible set of events, we introduce four probability functions, α, β, $\bar{\alpha}$, $\bar{\beta}$. The functions α and β characterize the simple situation, while $\bar{\alpha}$ and $\bar{\beta}$ characterize the canceling situation. First, for the simple situation, define the probability function $\alpha : \mathbb{N} \times \{0, \dots, q-1\} \times \{0, \dots, q-1\} \to [0, 1]$ as follows: given a group of $s_0 + s_1$ SDMs, $\alpha(t, s_0, s_1)$ is the probability that

(i) all the $s_0 + s_1$ SDMs have provided a decision up to time t; and
(ii) considering the variables $Count_0$ and $Count_1$ restricted to this group of $s_0 + s_1$ SDMs , $Count_0(t) = s_0$ and $Count_1(t) = s_1$.

Also, define the probability function $\beta_{1|j} : \mathbb{N} \times \{0, \dots, q-1\} \times \{0, \dots, q-1\} \to [0,1]$, $j \in \{0,1\}$ as follows: given a group of $N - (s_0 + s_1)$ SDMs, $\beta_{1|j}(t, s_0, s_1)$ is the probability that

(i) no SDMs have provided a decision up to time $t-1$; and
(ii) considering the variables $Count_0$ and $Count_1$ restricted to this group of $N-(s_0+s_1)$ SDMs, $Count_0(t) + s_0 < Count_1(t) + s_1$, and $Count_1(t) + s_1 \geq q$.

Similarly, it is straightforward to define the probabilities $\beta_{0|j}$, $j \in \{0,1\}$.

Second, for the canceling situation, define the probability function $\bar{\alpha} : \mathbb{N} \times \{q, \dots, \lfloor N/2 \rfloor\} \to [0,1]$ as follows: given a group of $2s$ SDMs, $\bar{\alpha}(t,s)$ is the probability that

(i) all the $2s$ SDMs have provided a decision up to time t; and
(ii) there exists $\bar{\tau} \leq t$ such that, considering the variables $Count_0$ and $Count_1$ restricted to this group of $2s$ SDMs
  - $Count_0(\bar{\tau}-1) < q$ and $Count_1(\bar{\tau}-1) < q$;
  - $Count_0(\tau) = Count_1(\tau) \geq q$ for all $\tau \geq \bar{\tau}$.

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

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

Similarly, it is straightforward to define the probabilities $\bar{\beta}_{0|j}$, $j \in \{0,1\}$.

Note that, for simplicity, we do not explicitly keep track of the dependence of the probabilities β and $\bar{\beta}$ upon the numbers N and q. The following proposition shows how to compute the probabilities $\left\{p_{i|j}(t; N, q)\right\}_{t=1}^{\infty}$, $i, j \in \{0,1\}$, starting from the above definitions.

Proposition III.1 ***(q out of N: a recursive formula)*** *Consider a group of N SDMs, running the q* out of *N* SDA *algorithm. Without loss of generality, assume H_1 is the correct hypothesis. Then, for $i \in \{0,1\}$, we have, for $t = 1$,*

$$p_{i|1}(1; N, q) = \beta_{i|1}(1, 0, 0), \tag{8}$$

and, for $t \geq 2$,

$$p_{i|1}(t; N, q) = \sum_{s_0=0}^{q-1} \sum_{s_1=0}^{q-1} \binom{N}{s_1 + s_0} \alpha(t-1, s_0, s_1) \beta_{i|1}(t, s_0, s_1) + \sum_{s=q}^{\lfloor N/2 \rfloor} \binom{N}{2s} \bar{\alpha}(t-1, s) \bar{\beta}_{i|1}(t, s). \tag{9}$$

Proof: The proof that formulas in (8) hold true follows trivially form the definition of the quantities $\beta_{1|1}(1,0,0)$ and $\beta_{0|1}(1,0,0)$. We start by providing three useful definitions.

First, let E_t denote the event that the SDA 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 SDMs that have decided in favor of H_0 up to time $t-1$;
(ii) there are s_1 SDMs 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 SDMs decides in favor of H_0 while r_1 SDMs 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) $2s$ SDMs 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 $2s$ SDMs
  - $Count_0(\tau) < q$, $Count_1(\tau) < q$, for $1 \leq \tau \leq \bar{\tau} - 1$;
  - $Count_0(\tau) = Count_1(\tau)$ for $\bar{\tau} \leq \tau \leq t-1$;
  - $Count_0(t-1) = Count_1(t-1) = s$.

(ii) at time t the number of SDMs providing their decision in favor of H_1 is strictly greater than the number of SDMs deciding in favor of H_0.

Observe that

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

Since $E_{s_0,s_1,t}$, $0 \le s_0, s_1 \le q-1$, and $E_{s,t}$, $q \le s \le \lfloor N/2 \rfloor$ are disjoint sets, we can write

$$\mathbb{P}[E_t] = \sum_{0 \le s_0, s_1 \le q-1} \mathbb{P}[E_{s_0,s_1,t}] + \sum_{q \le s \le \lfloor N/2 \rfloor} \mathbb{P}[E_{s,t}]. \tag{10}$$

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,

$$\mathbb{P}[E_{s_0,s_1,t}] = \binom{N}{s_1 + s_0} \alpha(t-1, s_0, s_1) \beta_{1|1}(t, s_0, s_1) \tag{11}$$

and that

$$\mathbb{P}[E_{s,t}] = \binom{N}{2s} \bar{\alpha}(t-1, s) \bar{\beta}_{1|1}(t, s). \tag{12}$$

Plugging equations (11) and (12) into equation (10) concludes the proof of the Theorem. ■

Formulas, similar to the ones in (8) and (9) can be provided for computing also the probabilities $\left\{ p_{i|0}(t; N, q) \right\}_{t=1}^{\infty}$, $i \in \{0, 1\}$.

As far as the probabilities $\alpha(t, s_0, s_1)$, $\bar{\alpha}(t, s)$, $\beta_{i|j}(t, s_0, s_1)$, $\bar{\beta}_{i|j}(t, s)$, $i, j \in \{0, 1\}$, are concerned, we now provide expressions to calculate them.

Proposition III.2 *Consider a group of N SDMs, running the q* out of N SDA *algorithm for $1 \le q \le \lfloor N/2 \rfloor$. Without loss of generality, assume H_1 is the correct hypothesis. For $i \in \{0, 1\}$, let $\pi_{i|1} : \mathbb{N} \to [0, 1]$ denote the cumulative probability up to time t that a single SDM provides the decision H_i, given that H_1 is the correct hypothesis, i.e.,*

$$\pi_{i|1}(t) = \sum_{s=1}^{t} p_{i|1}(t). \tag{13}$$

For $t \in \mathbb{N}$, $s_0, s_1 \in \{1, \dots, q-1\}$, $s \in \{q, \dots, \lfloor N/2 \rfloor\}$, the probabilities $\alpha(t, s_0, s_1)$, $\bar{\alpha}(t, s)$, $\beta_{1|1}(t, s_0, s_1)$, and $\bar{\beta}_{1|1}(t, s)$ satisfy the following relationships (explicit for α, β, $\bar{\beta}$ and recursive for $\bar{\alpha}$):

$$\alpha(t, s_0, s_1) = \binom{s_0 + s_1}{s_0} \pi_{0|1}^{s_0}(t) \pi_{1|1}^{s_1}(t),$$

$$\bar{\alpha}(t, s) = \sum_{s_0=0}^{q-1} \sum_{s_1=0}^{q-1} \binom{2s}{s_0 + s_1} \binom{2s - s_0 - s_1}{s - s_0} \alpha(t-1, s_0, s_1) p_{0|1}^{s-s_0}(t) p_{1|1}^{s-s_1}(t)$$
$$+ \sum_{h=q}^{s} \binom{2s}{2h} \binom{2s - 2h}{s - h} \bar{\alpha}(t-1, h) p_{0|1}^{s-h}(t) p_{1|1}^{s-h}(t),$$

$$\beta_{1|1}(t, s_0, s_1) = \sum_{h_1=q-s_1}^{N-\bar{s}} \binom{N - \bar{s}}{h_1} p_{1|1}^{h_1}(t) \left[\sum_{h_0=0}^{m} \binom{N - \bar{s} - h_1}{h_0} p_{0|1}^{h_0}(t) \left(1 - \pi_{1|1}(t) - \pi_{0|1}(t) \right)^{N - \bar{s} - h_0 - h_1} \right],$$

$$\bar{\beta}_{1|1}(t, s) = \sum_{h_1=1}^{N-2s} \binom{N - 2s}{h_1} p_{1|1}^{h_1}(t) \left[\sum_{h_0=0}^{\bar{m}} \binom{N - 2s - h_1}{h_0} p_{0|1}^{h_0}(t) (1 - \pi_{1|1}(t) - \pi_{0|1}(t))^{N - 2s - h_0 - h_1} \right],$$

where $\bar{s} = s_0 + s_1$, $m = \min\{h_1 + s_1 - s_0 - 1, N - (s_0 + s_1) - h_1\}$ and $\bar{m} = \min\{h_1 - 1, N - 2s - h_1\}$. Moreover, corresponding relationships 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 the relationships for $\beta_{1|1}(t, s_0, s_1)$ and $\bar{\beta}_{1|1}(t, s)$.

Proof: The evaluation of $\alpha(t, s_0, s_1)$ follows from standard probabilistic arguments. Indeed, observe that, given a first group of s_0 SDMs and a second group of s_1 SDMs, the probability that all the SDMs of the first group have decided in favor of H_0 up to time t and all the SDMs of the second group have decided in favor of

H_1 up to time t is given by $\pi_{0|1}^{s_0}(t)\pi_{1|1}^{s_1}(t)$. The desired result follows from the fact that there are $\binom{s_1+s_0}{s_0}$ ways of dividing a group of s_0+s_1 SDMs into two subgroups of s_0 and s_1 SDMs.

Consider now $\bar{\alpha}(t,s)$. Let $E_{\bar{\alpha}(t,s)}$ denote the event of which $\bar{\alpha}(t,s)$ is the probability of occurring, that is, the event that, given a group of $2s$ SDMs,

(i) all the $2s$ SDMs have provided a decision up to time t; and
(ii) there exists $\bar{\tau} \leq t$ such that, considering the variables $Count_0$ and $Count_1$ restricted to this group of $2s$ SDMs
  - $Count_0(\bar{\tau}-1) < q$ and $Count_1(\bar{\tau}-1) < q$;
  - $Count_0(\tau) = Count_1(\tau) \geq q$ for all $\tau \geq \bar{\tau}$.

Now, for a group of $2s$ SDMs, for $0 \leq s_0, s_1 \leq q-1$, let E_{t-1,s_0,s_1} denote the event that

(i) s_0 (resp. s_1) SDMs have decided in favor of H_0 (resp. H_1) up to time $t-1$;
(ii) $s-s_0$ (resp. $s-s_1$) SDMs decide in favor of H_0 (resp. H_1) at time t.

Observing that for s_0+s_1 assigned SDMs the probability that fact (i) is verified is given by $\alpha(t-1,s_0,s_1)$ we can write that

$$\mathbb{P}[E_{t-1,s_0,s_1}] = \binom{2s}{s_0+s_1}\binom{2s-s_0-s_1}{s-s_0}\alpha(t-1,s_0,s_1)p_{0|1}^{s-s_0}(t)p_{1|1}^{s-s_1}(t).$$

Consider again a group of $2s$ SDMs and for $q \leq h \leq s$ let $\bar{E}_{t-1,h}$ denote the event that

(i) $2h$ SDMs have provided a decision up to time $t-1$;
(ii) there exists $\bar{\tau} \leq t-1$ such that, considering the variables $Count_0$ and $Count_1$ restricted to the group of $2h$ SDMs that have already provided a decision,
  - $Count_0(\bar{\tau}-1) < q$ and $Count_1(\bar{\tau}-1) < q$;
  - $Count_0(\tau) = Count_1(\tau) \geq q$ for all $\tau \geq \bar{\tau}$; and
  - $Count_0(t-1) = Count_1(t-1) = h$;

(iii) at time instant t, $s-h$ SDMs decide in favor of H_0 and $s-h$ SDMs decide in favor of H_1.

Observing that for $2h$ assigned SDMs the probability that fact (i) and fact (ii) are verified is given by $\bar{\alpha}(t-1,h)$, we can write that

$$\mathbb{P}[\bar{E}_{t-1,h}] = \binom{2s}{2h}\binom{2s-2h}{s-h}\bar{\alpha}(t-1,h)p_{0|1}^{s-h}(t)p_{1|1}^{s-h}(t).$$

Observe that

$$E_{\bar{\alpha}(t,s)} = \left(\bigcup_{s_0=0}^{q}\bigcup_{s_1=0}^{q} E_{t-1,s_0,s_1}\right)\bigcup\left(\bigcup_{h=q}^{\lfloor N/2\rfloor}\bar{E}_{t-1,h}\right).$$

Since the events E_{t-1,s_0,s_1}, $0 \leq s_0, s_1 < q$ and $\bar{E}_{t-1,h}$, $q \leq h \leq \lfloor N/2 \rfloor$, are all disjoint we have that

$$\mathbb{P}[E_{\bar{\alpha}(t,s)}] = \sum_{s_0=0}^{q-1}\sum_{s_1=0}^{q-1}\mathbb{P}[E_{t-1,s_0,s_1}] + \sum_{h=q}^{s}\mathbb{P}[\bar{E}_{t-1,h}].$$

Plugging the expressions of $\mathbb{P}[E_{t-1,s_0,s_1}]$ and $\mathbb{P}[\bar{E}_{t-1,h}]$ in the above equality gives the recursive relationship for computing $\bar{\alpha}(t,s)$.

Consider now the probability $\beta_{1|1}(t,s_0,s_1)$. Recall that this probability refers to a group of $N-(s_0+s_1)$ SDMs. Let us introduce some notations. Let $E_{\beta_{1|1}(t,s_0,s_1)}$ denote the event of which $\beta_{1|1}(t,s_0,s_1)$ represents the probability of occurring and let $E_{t;h_1,s_1,h_0,s_0}$ denote the event that, at time t

- h_1 SDMs decides in favor of H_1;
- h_0 SDMs decides in favor of H_0;
- the remaining $N-(s_0+s_1)-(h_0+h_1)$ do not provide a decision up to time t.

Observe that the above event is well-defined if and only if $h_0+h_1 \leq N-(s_0+s_1)$. Moreover $E_{t;h_1,s_1,h_0,s_0}$ contributes to $\beta_{1|1}(t,s_0,s_1)$, i.e., $E_{t;h_1,s_1,h_0,s_0} \subseteq E_{\beta_{1|1}(t,s_0,s_1)}$ if and only if $h_1 \geq q-s_1$ and $h_0 < h_1+s_1-s_0$ (the necessity of these two inequalities follows directly from the definition of $\beta_{1|1}(t,s_0,s_1)$). Considering the three inequalities $h_0+h_1 \leq N-(s_0+s_1)$, $h_1 \geq q-s_1$ and $h_0 < h_1+s_1-s_0$, it follows that

$$E_{\beta_{1|1}(t,s_0,s_1)} = \bigcup\left\{E_{t;h_1,s_1,h_0,s_0} \mid q-s_1 \leq h_1 \leq N-(s_0+s_1) \text{ and } h_0 \leq m\right\},$$

where $m = \min\{h_1+s_1-s_0-1, N-(s_0+s_1)-h_1\}$. To conclude it suffices to observe that the events $E_{t;h_1,s_1,h_0,s_0}$ for $q-s_1 \le h_1 \le N-(s_0+s_1)$ and $h_0 \le m$ are disjoint events and that

$$\mathbb{P}[E_{t;h_1,s_1,h_0,s_0}] = \binom{N-\bar{s}}{j} p_{1|1}^{h_1}(t) \binom{N-\bar{s}-h_1}{h_0} p_{0|1}^{h_0}(t) \left(1-\pi_{1|1}(t)-\pi_{0|1}(t)\right)^{N-\bar{s}-h_0-h_1},$$

where $\bar{s} = s_0 + s_1$.

The probability $\bar{\beta}_{1|1}(t,s)$ can be computed reasoning similarly to $\beta_{1|1}(t,s_0,s_1)$. ■

Now we describe some properties of the above expressions in order to assess the computational complexity required by the formulas introduced in Proposition III.1 in order to compute $\left\{p_{i|j}(t;N,q)\right\}_{t=1}^{\infty}$, $i,j \in \{0,1\}$. From the expressions in Proposition III.2 we observe that

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

Moreover from equation (13) we have that $\pi_{i|j}(t)$ is a function of $\pi_{i|j}(t-1)$ 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 $\pi_{0|1}(t-1)$, $\pi_{1|1}(t-1)$, $\bar{\alpha}(t-1,s)$, $q \le h \le \lfloor N/2 \rfloor$ 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 $\left\{p_{i|j}(t;N,q)\right\}_{t=1}^{\infty}$, $i,j \in \{0,1\}$, requires keeping in memory a number of variables which grows linearly with the number of SDMs.

B. Case $\lfloor N/2 \rfloor + 1 \le q \le N$

The probabilities $p_{i|j}(t;N,q)$, $i,j \in \{0,1\}$ in the case where $\lfloor N/2\rfloor + 1 \le q \le N$ can be computed according to the expressions reported in the following Proposition.

Proposition III.3 *Consider a group of N SDMs, running the q* out of N SDA *algorithm for $\lfloor N/2\rfloor + 1 \le q \le N$. Without loss of generality, assume H_1 is the correct hypothesis. For $i \in \{0,1\}$, let $\pi_{i|1} : \mathbb{N} \to [0,1]$ be defined as* (13)*. Then, for $i \in \{0,1\}$, we have for $t=1$*

$$p_{i|1}(1;N,q) = \sum_{h=q}^{N} \binom{N}{h} p_{i|1}^{h}(1) \left(1-p_{i|1}(1)\right)^{N-h} \tag{14}$$

and for $t \ge 2$

$$p_{i|1}(t;N,q) = \sum_{k=0}^{q-1} \binom{N}{k} \pi_{i|1}^{k}(t-1) \sum_{h=q-k}^{N-k} \binom{N-k}{h} p_{i|1}^{h}(t) \left(1-\pi_{i|1}(t)\right)^{N-(h+k)}. \tag{15}$$

Proof: Let $t=1$. Since $q > N/2$, the probability that the fusion center decides in favor of H_i at time $t=1$ is given by the probability that al least q SDMs decide in favor of H_i at time 1. From standard combinatoric arguments this probability is given by (14).

If $t > 1$, the probability that the fusion center decides in favor of H_i at time t is given by the probability that h SDMs, $0 \le h < q$, have decided in favor of H_i up to time $t-1$, and that at least $q-h$ SDMs decide in favor of H_i at time t. Formally let $E_t^{(i)}$ denote the event that the fusion center provides its decision in favor of H_i at time t and let $E_{h,t;k,t-1}^{(i)}$ denote the event that k SDMs have decided in favor of H_i up to time $t-1$ and h SDMs decide in favor of H_i at time t. Observe that

$$E_t^{(i)} = \bigcup_{k=0}^{q-1} \bigcup_{h=q-k}^{N-k} E_{h,t;k,t-1}^{(i)}.$$

Since $E_{h,t;k,t-1}^{(i)}$ are disjoint sets it follows that

$$\mathbb{P}\left[E_t^{(i)}\right] = \sum_{k=0}^{q-1} \sum_{h=q-k}^{N-k} \mathbb{P}\left[E_{h,t;k,t-1}^{(i)}\right].$$

The proof is concluded by observing that

$$\mathbb{P}\left[E^{(i)}_{h,t;k,t-1}\right] = \binom{N}{k}\pi^k_{i|1}(t-1)\binom{N-k}{h}p^h_{i|1}(t)\left(1-\pi_{i|1}(t)\right)^{N-(h+k)}.$$

■

Regarding the complexity of the expressions in (15) it is easy to see that the probabilities $p_{i|j}(t;N,q)$, $i,j \in \{0,1\}$ can be computed as the output of a dynamical system having the two dimensional vector with components $\pi_{0|1}(t-1), \pi_{1|1}(t-1)$ as state and the two dimensional vector with components $p_{0|1}(t), p_{1|1}(t)$ as input. In this case the dimension of the system describing the evolution of the desired probabilities is independent of N.

IV. Scalability analysis of the fastest and majority sequential aggregation rules

The goal of this section is to provide some theoretical results characterizing the probabilities of being correct and wrong for a group implementing the *q-out-of-N* SDA rule. We also aim to characterize the probability with which such a group fails to reach a decision in addition to the time it takes for this group to stop running any test. In Sections IV-A and IV-B we consider the fastest and the majority rules, namely the thresholds $q=1$ and $q=\lceil N/2 \rceil$, respectively; we analyze how these two counting rules behave for increasing values of N. In Section IV-C, we study how these quantities vary with arbitrary values q and fixed values of N.

A. The fastest rule for varying values of N

In this section we provide interesting characterizations of accuracy and expected time under the *fastest* rule, i.e., the counting rules with threshold $q=1$. For simplicity we restrict to the case where the group has the *almost-sure* decision property. In particular we assume the following two properties.

Assumption IV.1 *The number N of SDMs is odd and the SDMs satisfy the* almost-sure *decision property.*

Here is the main result of this subsection. Recall that $p^{(\mathrm{f})}_{\mathrm{w}|1}(N)$ is the probability of wrong decision by a group of N SDMs implementing the fastest rule (assuming H_1 is the correct hypothesis).

Proposition IV.1 (Accuracy and expected time under the fastest rule) *Consider the* q out of N SDA *algorithm under Assumption IV.1. Assume $q=1$, that is, adopt the* fastest *SDA rule. Without loss of generality, assume H_1 is the correct hypothesis. Define the* earliest possible decision time

$$\bar{t} := \min\{t \in \mathbb{N} \mid \textit{either } p_{1|1}(t) \neq 0 \textit{ or } p_{0|1}(t) \neq 0\}. \tag{16}$$

Then the probability of error satisfies

$$\lim_{N\to\infty} p^{(\mathrm{f})}_{\mathrm{w}|1}(N) = \begin{cases} 0, & \textit{if } p_{1|1}(\bar{t}) > p_{0|1}(\bar{t}), \\ 1, & \textit{if } p_{1|1}(\bar{t}) < p_{0|1}(\bar{t}), \\ \frac{1}{2}, & \textit{if } p_{1|1}(\bar{t}) = p_{0|1}(\bar{t}), \end{cases} \tag{17}$$

and the expected decision time satisfies

$$\lim_{N\to\infty} \mathbb{E}\left[T|H_1, N, q=1\right] = \bar{t}. \tag{18}$$

Proof: We start by observing that in the case where the fastest rule is applied, formulas in (9) simplifies to

$$p_{1|1}(t;N,q=1) = \beta_{1|1}(t,0,0), \qquad \text{for all } t \in \mathbb{N}.$$

Now, since $p_{1|1}(t) = p_{0|1}(t) = 0$ for $t < \bar{t}$, it follows that

$$p_{1|1}(t;N,q=1) = \beta_{1|1}(t,0,0) = 0, \qquad t < \bar{t}.$$

Moreover we have $\pi_{1|1}(\bar{t}) = p_{1|1}(\bar{t})$ and $\pi_{0|1}(\bar{t}) = p_{0|1}(\bar{t})$. According to the definition of the probability $\beta_{1|1}(\bar{t},0,0)$, we write

$$\beta_{1|1}(\bar{t},0,0) = \sum_{j=1}^{N}\binom{N}{j}p^j_{1|1}(\bar{t})\left\{\sum_{i=0}^{m}\binom{N-j}{i}p^i_{0|1}(\bar{t})\left(1-p_{1|1}(\bar{t})-p_{0|1}(\bar{t})\right)^{N-i-j}\right\},$$

where $m = \min\{j-1, N-j\}$, or equivalently

$$\begin{aligned}\beta_{1|1}(\bar{t},0,0) = &\sum_{j=1}^{\lfloor N/2 \rfloor} \binom{N}{j} p_{1|1}^{j}(\bar{t}) \left\{ \sum_{i=0}^{j-1} \binom{N-j}{i} p_{0|1}^{i}(\bar{t}) \left(1 - p_{1|1}(\bar{t}) - p_{0|1}(\bar{t})\right)^{N-i-j} \right\} \\ &+ \sum_{j=\lceil N/2 \rceil}^{N} \binom{N}{j} p_{1|1}^{j}(\bar{t}) \left\{ \sum_{i=0}^{N-j} \binom{N-j}{i} p_{0|1}^{i}(\bar{t}) \left(1 - p_{1|1}(\bar{t}) - p_{0|1}(\bar{t})\right)^{N-i-j} \right\} \\ = &\sum_{j=1}^{\lfloor N/2 \rfloor} \binom{N}{j} p_{1|1}^{j}(\bar{t}) \left\{ \sum_{i=0}^{j-1} \binom{N-j}{i} p_{0|1}^{i}(\bar{t}) \left(1 - p_{1|1}(\bar{t}) - p_{0|1}(\bar{t})\right)^{N-i-j} \right\} \\ &+ \sum_{j=\lceil N/2 \rceil}^{N} \binom{N}{j} p_{1|1}^{j}(\bar{t}) \left(1 - p_{1|1}(\bar{t})\right)^{N-j}. \end{aligned} \tag{19}$$

An analogous expression for $\beta_{0|1}(\bar{t},0,0)$ can be obtained by exchanging the roles of $p_{0|1}(\bar{t})$ and $p_{0|1}(\bar{t})$ in equation (19). The rest of the proof is articulated as follows. First, we prove that

$$\lim_{N\to\infty} \left(p_{1|1}(\bar{t}; N, q=1) + p_{0|1}(\bar{t}; N, q=1)\right) = \lim_{N\to\infty} \left(\beta_{1|1}(\bar{t},0,0) + \beta_{0|1}(\bar{t},0,0)\right) = 1. \tag{20}$$

This fact implies that equation (18) holds and that, if $p_{1|1}(\bar{t}) = p_{0|1}(\bar{t})$, then $\lim_{N\to\infty} p_{\mathrm{w}|1}^{(\mathrm{f})}(N) = 1/2$. Indeed

$$\lim_{N\to\infty} \mathbb{E}\left[T|H_j, N, q=1\right] = \lim_{N\to\infty} \sum_{t=1}^{\infty} t(p_{0|j}(t; N, q=1) + p_{i|j}(t; N, q=1)) = \bar{t}.$$

Moreover, if $p_{1|1}(\bar{t}) = p_{0|1}(\bar{t})$, then also $(\beta_{1|1}(\bar{t},0,0) = \beta_{0|1}(\bar{t},0,0)$.

Second, we prove that $p_{1|1}(\bar{t}) > p_{0|1}(\bar{t})$ implies $\lim_{N\to\infty} \beta_{0|1}(\bar{t},0,0) = 0$. As a consequence, we have that $\lim_{N\to\infty} \beta_{1|1}(\bar{t},0,0) = 1$ or equivalently that $\lim_{N\to\infty} p_{\mathrm{w}|1}^{(\mathrm{f})}(N) = 0$.

To show equation (20), we consider the event *the group is not giving the decision at time* $\bar{t}$. We aim to show that the probability of this event goes to zero as $N \to \infty$. Indeed we have that

$$\mathbb{P}\left[T \neq \bar{t}\right] = \mathbb{P}\left[T > \bar{t}\right] = 1 - \left(p_{1|1}(\bar{t},N) + p_{0|1}(\bar{t},N)\right),$$

and, hence, $\mathbb{P}\left[T > \bar{t}\right] = 0$ implies $p_{1|1}(\bar{t},N) + p_{0|1}(\bar{t},N) = 1$. Observe that

$$\mathbb{P}\left[T > \bar{t}\right] = \sum_{j=0}^{\lfloor \frac{N}{2} \rfloor} \binom{N}{2j} \binom{2j}{j} p_{i|1}(\bar{t})^{j} p_{0|i}(\bar{t})^{j} \Big(1 - p_{i|1}(\bar{t}) - p_{0|i}(\bar{t})\Big)^{N-2j}.$$

For simplicity of notation, let us denote $x := p_{1|1}(\bar{t})$ and $y := p_{0|1}(\bar{t})$. We distinguish two cases, (i) $x \neq y$ and (ii) $x = y$.

Case $x \neq y$. We show that in this case there exists $\bar{\epsilon} > 0$, depending only on x and y, such that

$$\binom{2j}{j} x^{j} y^{j} < (x + y - \bar{\epsilon})^{2j}, \quad \text{for all } j \geq 1. \tag{21}$$

First of all observe that, since $\binom{2j}{j} x^j y^j$ is just one term of the Newton binomial expansion of $(x+y)^{2j}$, we know that $\binom{2j}{j} x^j y^j < (x+y)^{2j}$ for all $j \in \mathbb{N}$. Define $\epsilon(j) := x + y - \binom{2j}{j}^{1/2j} \sqrt{xy}$ and observe that proving equation (21) is equivalent to proving $\lim_{j\to\infty} \epsilon(j) > 0$. This also makes it possible to define $\bar{\epsilon} := \inf_{j\in\mathbb{N}} \epsilon(j)$. To prove the inequality $\lim_{j\to\infty} \epsilon(j) > 0$, let us compute $\lim_{j\to\infty} \binom{2j}{j}^{1/(2j)}$. By applying Stirling's formula we can write

$$\lim_{j\to\infty} \binom{2j}{j}^{1/(2j)} = \lim_{j\to\infty} \left(\frac{\sqrt{2\pi 2j} \left(\frac{2j}{e}\right)^{2j}}{2\pi j \left(\frac{j}{e}\right)^{2j}} \right)^{1/(2j)} = \left(\sqrt{\frac{1}{\pi j^2}} 2^{2j} \right)^{1/(2j)} = 2$$

and, in turn, $\lim_{j\to\infty}\epsilon(j) = x + y - 2\sqrt{xy}$. Clearly, if $x \neq y$, then $x + y - 2\sqrt{xy} > 0$. Defining $\bar{\epsilon} := \inf_{j\in\mathbb{N}}\epsilon(j)$, we can write

$$\begin{aligned}
\lim_{N\to\infty}\sum_{j=0}^{\lfloor \frac{N}{2}\rfloor}\binom{N}{2j}\binom{2j}{j}x^j y^j\,(1-x-y)^{N-2j} &\leq \lim_{N\to\infty}\sum_{j=0}^{\lfloor \frac{N}{2}\rfloor}\binom{N}{2j}(x+y-\bar{\epsilon})^{2j}\,(1-x-y)^{N-2j}\\
&\leq \lim_{N\to\infty}\sum_{j=0}^{N}\binom{N}{j}(x+y-\bar{\epsilon})^{j}\,(1-x-y)^{N-j}\\
&= \lim_{N\to\infty}(1-\bar{\epsilon})^N = 0,
\end{aligned}$$

which implies also $\lim_{N\to\infty}\mathbb{P}\left[T > \bar{t}\right] = 0$.

Case $x = y$. To study this case, let $y = x + \xi$ and let $\xi \to 0$. In this case, the probability of the decision time exceeding $\bar{t}$ becomes

$$f(x, N, \xi) = \mathbb{P}\left[T > \bar{t}\right] = \sum_{j=0}^{\lfloor \frac{N}{2}\rfloor}\binom{N}{2j}\binom{2j}{j}x^j (x+\xi)^j\,(1-2x-\xi)^{N-2j}.$$

Consider $\lim_{\xi\to 0} f(x, N, \xi)$. We have that

$$\lim_{\xi\to 0} f(x, N, \xi) = \sum_{j=0}^{\lfloor \frac{N}{2}\rfloor}\binom{N}{2j}\binom{2j}{j}x^{2j}\,(1-2x)^{N-2j} < \sum_{j=0}^{\lfloor \frac{N}{2}\rfloor}\binom{N}{2j}2^{2j}x^{2j}\,(1-2x)^{N-2j} < 1,$$

where the first inequality follows from $\binom{2j}{j} < \sum_{l=0}^{2j}\binom{2j}{l} = 2^{2j}$, and the second inequality follows from $\sum_{j=0}^{\lfloor \frac{N}{2}\rfloor}\binom{N}{2j}(2x)^{2j}\,(1-2x)^{N-2}$ $\sum_{2j=0}^{N}\binom{N}{2j}(2x)^{2j}\,(1-2x)^{N-2j} = 1$. So $\lim_{\xi\to 0} f(x, N, \xi)$ exists, and since we know that also $\lim_{N\to\infty} f(x, N, \xi)$ exists, the limits are exchangeable in $\lim_{N\to\infty}\lim_{\xi\to 0} f(x, N, \xi)$ and

$$\lim_{N\to\infty}\lim_{\xi\to 0} f(x, N, \epsilon) = \lim_{\xi\to 0}\lim_{N\to\infty} f(x, N, \xi) = 0.$$

This concludes the proof of equation (20).

Assume now that $p_{1|1}(\bar{t}) > p_{0|1}(\bar{t})$. We distinguish between the case where $p_{1|1}(\bar{t}) > \frac{1}{2}$ and the case where $p_{0|1}(\bar{t}) < p_{1|1}(\bar{t}) \leq \frac{1}{2}$.

If $p_{1|1}(\bar{t}) > \frac{1}{2}$, then Lemma .1 implies

$$\lim_{N\to\infty}\sum_{j=\lceil N/2\rceil}^{N}\binom{N}{j}p_{1|1}^j(\bar{t})\left(1-p_{1|1}(\bar{t})\right)^{N-j} = 1,$$

and, since $\lim_{N\to\infty}\beta_{1|1}(\bar{t}, 0, 0) > \lim_{N\to\infty}\sum_{j=\lceil N/2\rceil}^{N}\binom{N}{j}p_{1|1}^j(\bar{t})\left(1-p_{1|1}(\bar{t})\right)^{N-j}$, we have also that $\lim_{N\to\infty}\beta_{1|1}(\bar{t}, 0, 0) = 1$.

The case $p_{0|1}(\bar{t}) < p_{1|1}(\bar{t}) < \frac{1}{2}$ is more involved. We will see that in this case $\lim_{N\to\infty}\beta_{0|1}(\bar{t}, 0, 0) = 0$. We start by observing that, from Lemma .1,

$$\lim_{N\to\infty}\sum_{j=\lceil \frac{N}{2}\rceil}^{N}\binom{N}{j}p_{1|1}^j(\bar{t})\left([1-p_{1|1}(\bar{t})\right)^{N-j} = 0,$$

and in turn

$$\lim_{N\to\infty}\beta_{1|1}(\bar{t}, 0, 0) = \lim_{N\to\infty}\sum_{j=1}^{\lfloor \frac{N}{2}\rfloor}\binom{N}{j}p_{1|1}^j(\bar{t}) \times \left(\sum_{i=0}^{j-1}\binom{N-j}{i}p_{0|1}^i(\bar{t})\left[1-p_{1|1}(\bar{t})-p_{0|1}(\bar{t})\right]^{N-j-i}\right).$$

The above expression can be written as follows

$$\lim_{N\to\infty}\beta_{1|1}(\bar{t},0,0) = \lim_{N\to\infty}\sum_{h=1}^{N-2}\Bigg(\sum_{j=\lfloor\frac{h}{2}\rfloor+1}^{h}\binom{N}{j}\binom{N-j}{h-j}p_{0|1}^{h-j}(\bar{t})p_{1|1}^{j}(\bar{t})\Bigg)\Bigg(1-\big(p_{0|1}(\bar{t})p_{1|1}(\bar{t})\big)\Bigg)^{N-h}$$
$$= \lim_{N\to\infty}\sum_{h=1}^{N-2}\binom{N}{h}\sum_{j=\lfloor\frac{h}{2}\rfloor+1}^{h}\binom{h}{j}p_{1|1}^{h-j}(\bar{t})p_{0|1}^{j}(\bar{t})\Bigg(1-p_{1|1}(\bar{t})-p_{0|1}(\bar{t})\Bigg)^{N-h}$$

where, for obtaining the second equality we used the fact $\binom{N}{j}\binom{N-j}{h-j}=\binom{N}{h}\binom{h}{j}$. Similarly,

$$\lim_{N\to\infty}\beta_{0|1}(\bar{t},0,0) = \lim_{N\to\infty}\sum_{h=1}^{N-2}\binom{N}{h}\sum_{j=\lfloor\frac{h}{2}\rfloor+1}^{h}\binom{h}{j}p_{0|1}^{h-j}(\bar{t})p_{1|1}^{j}(\bar{t})\Bigg(1-p_{1|1}(\bar{t})-p_{0|1}(\bar{t})\Bigg)^{N-h}.$$

We prove now that $\lim_{N\to\infty}\beta_{0|1}(\bar{t},0,0)=0$. To do so we will show that there exists $\bar{\epsilon}$ depending only on $p_{0|1}(\bar{t})$ and $p_{1|1}(\bar{t})$ such that

$$\sum_{j=\lfloor\frac{h}{2}\rfloor+1}^{h}\binom{h}{j}p_{0|1}^{h-j}(\bar{t})p_{1|1}^{j}(\bar{t}) < \Bigg(p_{0|1}(\bar{t})+p_{1|1}(\bar{t})-\bar{\epsilon}\Bigg)^{h}.$$

To do so, let

$$\epsilon(h) = p_{0|1}(\bar{t})+p_{1|1}(\bar{t})-\sqrt[h]{\sum_{j=\lfloor\frac{h}{2}\rfloor+1}^{h}\binom{h}{j}p_{0|1}^{h-j}(\bar{t})p_{1|1}^{j}(\bar{t})}.$$

Because h is bounded, one can see that $\epsilon(h)>0$ as the sum inside the root is always smaller than $(p_{0|1}(\bar{t})+p_{1|1}(\bar{t}))^h$. We show below that Lemma .1 implies that $\liminf_{h\to\infty}\epsilon(h)>0$. In fact, with the notation in Lemma .1 and with $x=p_{0|1}(\bar{t})$ and $c=p_{0|1}(\bar{t})+p_{1|1}(\bar{t})$,

$$\sum_{j=\lfloor\frac{h}{2}\rfloor+1}^{h}\binom{h}{j}p_{0|1}^{h-j}(\bar{t})p_{1|1}^{j}(\bar{t}) = \sum_{j=\lfloor\frac{h}{2}\rfloor+1}^{h}\binom{h}{j}x^{h-j}(c-x)^{j} = \overline{S}(h;c,x),$$

and, using Lemma .1,

$$\frac{\overline{S}(h;c,x)}{c^h} < \frac{\lceil h/2\rceil\binom{h}{\lceil h/2\rceil}x^{\lceil h/2\rceil}(c-x)^{\lfloor h/2\rfloor}}{c^h}.$$

From the above, we know that

$$\epsilon(h) > c-\sqrt[h]{\lceil h/2\rceil\binom{h}{\lceil h/2\rceil}x^{\lceil h/2\rceil}(c-x)^{\lfloor h/2\rfloor}}.$$

We now take $\liminf_{h\to\infty}\epsilon(h)$ and obtain

$$\liminf_{h\to\infty}\epsilon(h) \ge c-\lim_{h\to\infty}\lceil h/2\rceil^{1/h}\binom{h}{\lceil h/2\rceil}^{1/h}(x^{\lceil h/2\rceil})^{1/h}(c-x)^{\lfloor h/2\rfloor/h} = c-2x^{1/2}(c-x)^{1/2}.$$

The last equality is true because $\lim_{h\to\infty}\lceil h/2\rceil^{1/h}\binom{h}{\lceil h/2\rceil}^{1/h}=2$. In order to complete the proof, we need to show that $c-2x^{1/2}(c-x)^{1/2}>0$, indeed $c-2x^{1/2}(c-x)^{1/2}$ has the same sign as $c^2-4x(c-x)=(c-2x)^2$. Since we are studying the case where $p_{0|1}(\bar{t})<p_{1|1}(\bar{t})$, it follows that $0\le x<\frac{c}{2}$ making the term of interest always strictly positive. This fact implies that $\liminf_{h\to\infty}\epsilon(h)$ is strictly positive and that, along with the fact that $\epsilon(h)>0$ for all finite h by definition, $\inf_{h\in\mathbb{N}}\epsilon(h)>0$.

By letting $\bar{\epsilon} := \inf_{h\in\mathbb{N}} \epsilon(h)$, we conclude that

$$\lim_{N\to\infty} \beta_{0|1}(\bar{t},0,0) \leq \sum_{h=1}^{N-2} \binom{N}{h} \Big(p_{1|1}(\bar{t}) + p_{0|1}(\bar{t}) - \bar{\epsilon}\Big)\Big(1 - p_{1|1}(\bar{t}) - p_{0|1}(\bar{t})\Big)^{N-h}$$
$$\leq \sum_{h=0}^{N} \binom{N}{h} \Big(p_{1|1}(\bar{t}) + p_{0|1}(\bar{t}) - \bar{\epsilon}\Big)\Big(1 - p_{1|1}(\bar{t}) - p_{0|1}(\bar{t})\Big)^{N-h} = (1-\bar{\epsilon})^N = 0.$$

This concludes the proof. ■

Remark IV.2 The earliest possible decision time $\bar{t}$ defined in (16) is the best performance that the fastest rule can achieve in terms of number of iterations required to provide the final decision. □

B. The majority rule for varying values of N

We consider now the *majority* rule, i.e., the counting rule with threshold $q = \lfloor N/2 \rfloor + 1$. We start with the following result about the accuracy. Recall that $p_{\mathrm{w}|1}$ is the probability of wrong decision by a single SDM and that $p_{\mathrm{w}|1}^{(\mathrm{m})}(N)$ is the probability of wrong decision by a group of N SDMs implementing the majority rule (assuming H_1 is the correct hypothesis).

Proposition IV.3 (Accuracy under the majority rule) *Consider the q out of N SDA algorithm under Assumption IV.1. Assume $q = \lfloor N/2 \rfloor + 1$, i.e., the* majority *rule is adopted. Without loss of generality, assume H_1 is the correct hypothesis. Then the probability of error satisfies*

$$p_{\mathrm{w}|1}^{(\mathrm{m})}(N) = \sum_{j=\lfloor N/2 \rfloor+1}^{N} \binom{N}{j} p_{\mathrm{w}|1}^{j} \left(1 - p_{\mathrm{w}|1}\right)^{N-j}. \qquad (22)$$

According to (22)*, the following characterization follows:*

(i) *if $0 \leq p_{\mathrm{w}|1} < 1/2$, then $p_{\mathrm{w}|1}^{(\mathrm{m})}(N)$ is a monotonic decreasing function of N that approaches 0 asymptotically, that is,*

$$p_{\mathrm{w}|1}^{(\mathrm{m})}(N) > p_{\mathrm{w}|1}^{(\mathrm{m})}(N+2) \quad \textit{and} \quad \lim_{N\to\infty} p_{\mathrm{w}|1}^{(\mathrm{m})}(N) = 0;$$

(ii) *if $1/2 < p_{\mathrm{w}|1} \leq 1$, then $p_{\mathrm{w}|1}^{(\mathrm{m})}(N)$ is a monotonic increasing function of N that approaches 1 asymptotically, that is,*

$$p_{\mathrm{w}|1}^{(\mathrm{m})}(N) < p_{\mathrm{w}|1}^{(\mathrm{m})}(N+2) \quad \textit{and} \quad \lim_{N\to\infty} p_{\mathrm{w}|1}^{(\mathrm{m})}(N) = 1;$$

(iii) *if $p_{\mathrm{w}|1} = 1/2$, then $p_{\mathrm{w}|1}^{(\mathrm{m})}(N) = 1/2$;*

(iv) *if $p_{\mathrm{w}|1} < 1/4$, then*

$$p_{\mathrm{w}|1}^{(\mathrm{m})}(N) = \binom{N}{\lceil \frac{N}{2} \rceil} p_{\mathrm{w}|1}^{\lceil \frac{N}{2} \rceil} + o\left(p_{\mathrm{w}|1}^{\lceil \frac{N}{2} \rceil}\right) = \sqrt{N/(2\pi)}\,(4p_{\mathrm{w}|1})^{\lceil \frac{N}{2} \rceil} + o\left((4p_{\mathrm{w}|1})^{\lceil \frac{N}{2} \rceil}\right). \qquad (23)$$

Proof: We start by observing that

$$\sum_{s=1}^{t} p_{0|1}(s; N, q = \lfloor N/2 \rfloor + 1) = \sum_{j=\lfloor N/2 \rfloor+1}^{N} \binom{N}{j} \pi_{0|1}(t)^j \left(1 - \pi_{0|1}(t)\right)^{N-j}.$$

Since $p_{\mathrm{w}|1}^{(\mathrm{m})}(N) = \sum_{s=1}^{\infty} p_{0|1}(s; N, q = \lfloor N/2 \rfloor + 1)$, taking the limit for $t \to \infty$ in the above expression leads to

$$p_{\mathrm{w}|1}^{(\mathrm{m})}(N) = \sum_{j=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{j} p_{\mathrm{w}|1}^{j} \left(1 - p_{\mathrm{w}|1}\right)^{N-j}.$$

Facts (i), (ii), (iii) follow directly from Lemma .1 in Appendix A applied to equation (22). Equation (23) is a consequence of the Taylor expansion of (22):

$$\begin{aligned}\sum_{j=\lceil \frac{N}{2}\rceil}^{N}\binom{N}{j}p_{\mathrm{w}|1}^{j}(1-p_{\mathrm{w}|1})^{N-j} &= \sum_{j=\lceil \frac{N}{2}\rceil}^{N}\binom{N}{j}p_{\mathrm{w}|1}^{j}(1-(N-j)p_{\mathrm{w}|1}+o(p_{\mathrm{w}|1}))\\ &= \binom{N}{\lceil \frac{N}{2}\rceil}p_{\mathrm{w}|1}^{\lceil \frac{N}{2}\rceil}+o\left(p_{\mathrm{w}|1}^{\lceil \frac{N}{2}\rceil+1}\right).\end{aligned}$$

Finally, Stirling's Formula implies $\lim_{N\to\infty}\binom{N}{\lceil \frac{N}{2}\rceil}=\sqrt{2N/\pi}\,2^N$ and, in turn, the final expansion follows from $2^N=4^{\lceil N/2\rceil}/2$. ∎

We discuss now the expected time required by the collective SDA algorithm to provide a decision when the *majority* rule is adopted. Our analysis is based again on Assumption IV.1 and on the assumption that H_1 is the correct hypothesis. We distinguish four cases based on different properties that the probabilities of wrong and correct decision of the single SDM might have:

(A1) the probability of correct decision is greater than the probability of wrong decision, i.e., $p_{\mathrm{c}|1}>p_{\mathrm{w}|1}$;

(A2) the probability of correct decision is equal to the probability of wrong decision, i.e., $p_{\mathrm{c}|1}=p_{\mathrm{w}|1}=1/2$ and there exist t_0 and t_1 such that $\pi_{0|1}(t_0)=1/2$ and $\pi_{1|1}(t_1)=1/2$;

(A3) the probability of correct decision is equal to the probability of wrong decision, i.e., $p_{\mathrm{c}|1}=p_{\mathrm{w}|1}=1/2$ and there exists t_1 such that $\pi_{1|1}(t_1)=1/2$, while $\pi_{0|1}(t)<1/2$ for all $t\in\mathbb{N}$ and $\lim_{t\to\infty}\pi_{0|1}(t)=1/2$;

(A4) the probability of correct decision is equal to the probability of wrong decision, i.e., $p_{\mathrm{c}|1}=p_{\mathrm{w}|1}=1/2$, and $\pi_{0|1}(t)<1/2$, $\pi_{1|1}(t)<1/2$ for all $t\in\mathbb{N}$ and $\lim_{t\to\infty}\pi_{0|1}=\lim_{t\to\infty}\pi_{1|1}(t)=1/2$.

Note that, since Assumption IV.1 implies $p_{\mathrm{c}|1}+p_{\mathrm{w}|1}=1$, the probability of correct decision in case (A1) satisfies $p_{\mathrm{c}|1}>1/2$. Hence, in case (A1) and under Assumption IV.1, we define $t_{<\frac{1}{2}}:=\max\{t\in\mathbb{N}\mid \pi_{1|1}(t)<1/2\}$ and $t_{>\frac{1}{2}}:=\min\{t\in\mathbb{N}\mid \pi_{1|1}(t)>1/2\}$.

Proposition IV.4 (Expected time under the *majority* rule) *Consider the q out of N SDA algorithm under Assumption IV.1. Assume $q=\lfloor N/2\rfloor+1$, that is, adopt the* majority *rule. Without loss of generality, assume H_1 is the correct hypothesis. Define the SDM properties (A1)-(A4) and the decision times t_0, t_1, $t_{<\frac{1}{2}}$ and $t_{>\frac{1}{2}}$ as above. Then the expected decision time satisfies*

$$\lim_{N\to\infty}\mathbb{E}\big[T|H_1,N,q=\lceil N/2\rceil\big]=\begin{cases}\dfrac{t_{<\frac{1}{2}}+t_{>\frac{1}{2}}+1}{2}, & \textit{if the SDM has the property (A1)},\\ \dfrac{t_1+t_0}{2}, & \textit{if the SDM has the property (A2)},\\ +\infty, & \textit{if the SDM has the property (A3) or (A4)}.\end{cases}$$

Proof: We start by proving the equality for case (A1). Since, in this case we are assuming $p_{\mathrm{c}|1}>p_{\mathrm{w}|1}$, the definitions of $t_{<\frac{1}{2}}$ and $t_{>\frac{1}{2}}$ implies that $\pi_{1|1}(t)=1/2$ for all $t_{<\frac{1}{2}}<t<t_{>\frac{1}{2}}$. Observe that

$$\sum_{s=1}^{t}p_{1|1}(s;N,q=\lfloor N/2\rfloor+1)=\sum_{h=\lfloor\frac{N}{2}\rfloor}^{N}\binom{N}{h}\pi_{1|1}^{h}(t)\Big(1-\pi_{1|1}(t)\Big)^{N-h}.$$

Hence Lemma .1 implies

$$\lim_{N\to\infty}\sum_{s=1}^{t}p_{1|1}(s;N,q=\lfloor N/2\rfloor+1)=\begin{cases}0, & \text{if } t\le t_{<\frac{1}{2}},\\ 1, & \text{if } t\ge t_{>\frac{1}{2}},\\ \frac{1}{2}, & \text{if } t_{<\frac{1}{2}}<t<t_{>\frac{1}{2}},\end{cases}$$

and, in turn, that

$$\lim_{N\to\infty}p_{1|1}(t;N,q=\lfloor N/2\rfloor+1)=\begin{cases}1/2, & \text{if } t=t_{<\frac{1}{2}}+1 \text{ and } t=t_{>\frac{1}{2}},\\ 0, & \text{otherwise}.\end{cases}$$

It follows

$$\lim_{N\to\infty} \mathbb{E}\left[T|H_1, N, q = \lfloor N/2 \rfloor + 1\right] = \lim_{N\to\infty} \sum_{t=0}^{\infty} t \left(p_{0|1}(t; N, q = \lfloor N/2 \rfloor + 1) + p_{1|1}(t; N, q = \lfloor N/2 \rfloor + 1)\right)$$
$$= \frac{1}{2}\left(t_{<\frac{1}{2}} + 1 + t_{>\frac{1}{2}}\right).$$

This concludes the proof of the equality for case (A1).

We consider now the case (A2). Reasoning similarly to the previous case we have that

$$\lim_{N\to\infty} p_{1|1}(t_1; N, q = \lfloor N/2 \rfloor + 1) = 1/2 \qquad \text{and} \qquad \lim_{N\to\infty} p_{0|1}(t_0; N, q = \lfloor N/2 \rfloor + 1) = 1/2,$$

from which it easily follows that $\lim_{N\to\infty} \mathbb{E}\left[T|H_1, N, q = \lfloor N/2 \rfloor + 1\right] = \frac{1}{2}(t_0 + t_1)$.

For case (A3), it suffices to note the following implication of Lemma .1: if, for a given $i \in \{0, 1\}$, we have $\pi_{i|1}(t) < 1/2$ for all $t \in \mathbb{N}$, then $\lim_{N\to\infty} p_{i|1}(t; N, q = \lfloor N/2 \rfloor + 1) = 0$ for all $t \in \mathbb{N}$. The analysis of the case (A4) is analogous to that of case (A3). ■

Remark IV.5 The cases where $p_{\mathrm{w}|1} > p_{\mathrm{c}|1}$ and where there exists t_0 such that $\pi_{0|1}(t_0) = 1/2$ while $\pi_{1|1}(t) < 1/2$ for all $t \in \mathbb{N}$ and $\lim_{t\to\infty} \pi_{1|1}(t) = 1/2$, can be analyzed similarly to the cases (A1) and (A3). Moreover, the most recurrent situation in applications is the one where there exists a time instant t such that $\pi_{1|1}(t) < 1/2$ and $\pi_{1|1}(t + 1) > 1/2$, which is equivalent to the above case (A1) with $t_{>\frac{1}{2}} = t_{<\frac{1}{2}} + 1$. In this situation we trivially have $\lim_{N\to\infty} \mathbb{E}\left[T|H_1, N, q = \lceil N/2 \rceil\right] = t_{>\frac{1}{2}}$. □

C. *Fixed N and varying q*

We start with a simple result characterizing the expected decision time.

Proposition IV.6 *Given a group of N SDMs running the q* out of N SDA, *for* $j \in \{0, 1\}$,

$$\mathbb{E}[T|H_j, N, q = 1] \leq \mathbb{E}[T|H_j, N, q = 2] \leq \cdots \leq \mathbb{E}[T|H_j, N, q = N].$$

The above proposition states that the expected number of iterations required to provide a decision constitutes a nondecreasing sequence for increasing value of q. Similar monotonicity results hold true also for $p_{\mathrm{c}|j}(N, q)$, $p_{\mathrm{w}|j}(N, q)$, $p_{\mathrm{nd}|j}(N, q)$ even though restricted only to $\lfloor N/2 \rfloor + 1 \leq q \leq N$.

Proposition IV.7 *Given a group of N SDMs running the q* out of N SDA, *for* $j \in \{0, 1\}$,

$$p_{\mathrm{c}|j}(N, q = \lfloor N/2 \rfloor + 1) \geq p_{\mathrm{c}|j}(N, q = \lfloor N/2 \rfloor + 2) \geq \cdots \geq p_{\mathrm{c}|j}(N, q = N),$$
$$p_{\mathrm{w}|j}(N, q = \lfloor N/2 \rfloor + 1) \geq p_{\mathrm{w}|j}(N, q = \lfloor N/2 \rfloor + 2) \geq \cdots \geq p_{\mathrm{w}|j}(N, q = N),$$
$$p_{\mathrm{nd}|j}(N, q = \lfloor N/2 \rfloor + 1) \leq p_{\mathrm{nd}|j}(N, q = \lfloor N/2 \rfloor + 2) \leq \cdots \leq p_{\mathrm{nd}|j}(N, q = N).$$

We believe that similar monotonic results hold true also for $1 \leq q \leq \lfloor N/2 \rfloor$. In particular, here is our conjecture: if N is odd, the single SDM has the *almost-sure* decision and the single SDM is more likely to provide the correct decision than the wrong decision, that is, $p_{\mathrm{c}|j} + p_{\mathrm{w}|j} = 1$ and $p_{\mathrm{c}|j} > p_{\mathrm{w}|j}$, then

$$p_{\mathrm{c}|j}(N, q = 1) \leq p_{\mathrm{c}|j}(N, q = 2) \leq \cdots \leq p_{\mathrm{c}|j}(N, q = \lfloor N/2 \rfloor + 1),$$
$$p_{\mathrm{w}|j}(N, q = 1) \geq p_{\mathrm{w}|j}(N, q = 2) \geq \cdots \geq p_{\mathrm{w}|j}(N, q = \lfloor N/2 \rfloor + 1).$$

These chains of inequalities are numerically verified in some examples in Section V.

V. NUMERICAL ANALYSIS

The goal of this section is to numerically analyze the models and methods described in previous sections. In all the examples, we assume that the sequential binary test run by each SDMs is the classical sequential probability ratio test (SPRT) developed in 1943 by Abraham Wald. To fix some notation, we start by briefly reviewing the SPRT. Let X be a random variable with distribution $f(x;\theta)$ and assume the goal is to test the null hypothesis $H_0 : \theta = \theta_0$ against the alternative hypothesis $H_1 : \theta = \theta_1$. For $i \in \{1, \dots, N\}$, the i-th SDM takes the observations $x_i(1), x_i(2), x(3), \dots$, which are assumed to be independent of each other and from the observations taken by all the other SDMs. The log-likelihood ratio associated to the observation $x_i(t)$ is

$$\lambda_i(t) = \log \frac{f(x_i(t), \theta_1)}{f(x_i(t), \theta_0)}. \tag{24}$$

Accordingly, let $\Lambda_i(t) = \sum_{h=1}^{t} \lambda_i(h)$ denote the sum of the log-likelihoods up to time instant t. The i-th SDM continues to sample as long as $\eta_0 < \Lambda_i(t) < \eta_1$, where η_0 and η_1 are two pre-assigned thresholds; instead sampling is stopped the first time this inequality is violated. If $\Lambda_i(t) < \eta_0$, then the i-th SDM decides for $\theta = \theta_0$. If $\Lambda_i(t) > \eta_1$, then the i-th SDM decides for $\theta = \theta_1$.

To guarantee the *homogeneity property* we assume that all the SDMs have the same thresholds η_0 and η_1. The threshold values are related to the accuracy of the SPRT as described in the classic Wald's method [19]. We shortly review this method next. Assume that, for the single SDM, we want 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 $p_{\text{misdetection}}$ and $p_{\text{false alarm}}$. Wald proved that the inequalities $\mathbb{P}[\text{say } H_0 \mid H_1] \leq p_{\text{misdetection}}$ and $\mathbb{P}[\text{say } H_1 \mid H_0] \leq p_{\text{false alarm}}$ are achieved when η_0 and η_1 satisfy $\eta_0 \leq \log \frac{p_{\text{misdetection}}}{1-p_{\text{false alarm}}}$ and $\eta_1 \geq \log \frac{1-p_{\text{misdetection}}}{p_{\text{false alarm}}}$. As customary, we adopt the equality sign in these inequalities for the design of η_0 and η_1. Specifically, in all our examples we assume that $p_{\text{misdetection}} = p_{\text{false alarm}} = 0.1$ and, in turn, that $\eta_1 = -\eta_0 = \log 9$.

We provide numerical results for observations described by both discrete and continuous random variables. In case X is a discrete random variable, we assume that $f(x;\theta)$ is a binomial distribution

$$f(x;\theta) = \begin{cases} \binom{n}{x} \theta^x (1-\theta)^{n-x}, & \text{if } x \in \{0, 1, \dots, n\}, \\ 0, & \text{otherwise,} \end{cases} \tag{25}$$

where n is a positive integer number. In case X is a continuous random variable, we assume that $f(x;\theta)$ is a Gaussian distribution with mean θ and variance σ^2

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\theta)^2/2\sigma^2}. \tag{26}$$

The key ingredient required for the applicability of Propositions III.1 and III.2 is the knowledge of the probabilities $\{p_{0|0}(t), p_{1|0}(t)\}_{t\in\mathbb{N}}$ and $\{p_{0|1}(t), p_{1|1}(t)\}_{t\in\mathbb{N}}$. Given thresholds η_0 and η_1, there probabilities can be computed according to the method described in the Appendix B.1 (respectively Appendix B.2) for X discrete (respectively X continuous) random variable.

We provide three sets of numerical results. Specifically, in Example V.1 we emphasize the tradeoff between accuracy and expected decision time as a function of the number of SDMs. In Example V.2 we concentrate on the monotonic behaviors that the q *out of* N SDA algorithm exhibits both when N is fixed and q varies and when q is fixed and N varies. In Example V.3 we compare the *fastest* and the *majority* rule. Finally, Section V-A discusses drawing connections between the observations in Example V.3 and the cognitive psychology presentation introduced in Section I-D.

Example V.1 (Tradeoff between accuracy and expected decision time) This example emphasizes the tradeoff between accuracy and expected decision time as a function of the number of SDMs. We do that for the *fastest* and the *majority* rules. We obtain our numerical results for odd sizes of group of SDMs ranging from 1 to 61. In all our numerical examples, we compute the values of the thresholds η_0 and η_1 according to Wald's method by posing $p_{\text{misdetection}} = p_{\text{false alarm}} = 0.1$ and, therefore, $\eta_1 = \log 9$ and $\eta_0 = -\log 9$.

For a binomial distribution $f(x;\theta)$ as in (25), we provide our numerical results under the following conditions: we set $n = 5$; we run our computations for three different pairs (θ_0, θ_1); precisely we assume that $\theta_0 = 0.5 - \epsilon$ and

$\theta_1 = 0.5 + \epsilon$ where $\epsilon \in \{0.03, 0.05, 0.07\}$; and $H_1 : \theta = \theta_1$ is always the correct hypothesis. For any pair (θ_0, θ_1) we perform the following three actions in order

(i) we compute the probabilities $\{p_{0|1}(t), p_{1|1}(t)\}_{t\in\mathbb{N}}$ according to the method described in Appendix B.1;

(ii) we compute the probabilities $\{p_{0|1}(t; N, q), p_{1|1}(t; N, q)\}_{t\in\mathbb{N}}$ for $q = 1$ and $q = \lfloor N/2 \rfloor + 1$ according to the formulas reported in Proposition III.1;

(iii) we compute probability of wrong decision and expected time for the group of SDMs exploiting the formulas

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

According to Remark II.7, since we consider only odd numbers N of SDMs, since $q \leq \lceil N/2 \rceil$ and since each SDM running the SPRT has the *almost-sure decisions* property, then $p_{\mathrm{w}|1}(N, q) + p_{\mathrm{c}|1}(N, q) = 1$. In other words, the probability of no-decision is equal to 0 and, hence, the accuracy of the SDA algorithms is characterized only by the probability of wrong decision and the probability of correct decision. In our analysis we select to compute the probability of wrong decision.

For a Gaussian distribution $f(x; \theta, \sigma)$, we obtain our numerical results under the following conditions: the two hypothesis are $H_0 : \theta = 0$ and $H_1 : \theta = 1$; we run our computations for three different values of σ; precisely $\sigma \in \{0.5, 1, 2\}$; and $H_1 : \theta = 1$ is always the correct hypothesis. To obtain $p_{\mathrm{w}|1}(N, q)$ and $\mathbb{E}[T|H_1, N, q]$ for a given value of σ, we proceed similarly to the previous case with the only difference that $\{p_{0|1}(t), p_{1|1}(t)\}_{t\in\mathbb{N}}$ are computed according to the procedure described in Appendix B.2.

The results obtained for the *fastest* rule are depicted in Figure V.1, while the results obtained for the *majority* rule are reported in Figure 3.

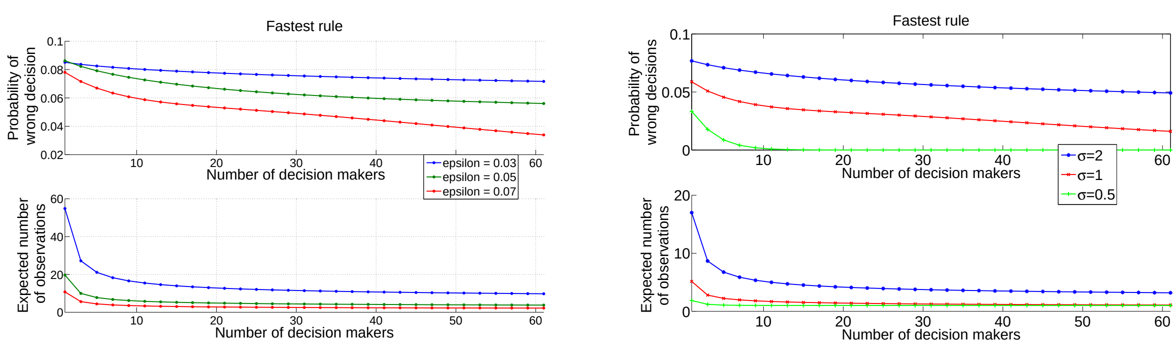

Fig. 2. Behavior of the probability of wrong decision and of the expected number of iterations required to provide a decision as the number of SDMs increases when the *fastest* rule is adopted. In Figure (a) we consider the binomial distribution, in Figure (b) the Gaussian distribution.

Some remarks are now in order. We start with the *fastest* rule. A better understanding of the plots in Figure V.1 can be gained by specifying the values of the earliest possible decision time $\bar{t}$ defined in (16) and of the probabilities $p_{1|1}(\bar{t})$ and $p_{0|1}(\bar{t})$. In our numerical analysis, for each pair (θ_0, θ_1) considered and for both discrete and continuous measurements X, we had $\bar{t} = 1$ and $p_{1|1}(\bar{t}) > p_{0|1}(\bar{t})$. As expected from Proposition IV.1, we can see that the *fastest* rule significantly reduces the expected number of iterations required to provide a decision. Indeed, as N increases, the expected decision time $\mathbb{E}[T|H_1, N, q = 1]$ tends to 1. Moreover, notice that $p^{(\mathrm{f})}_{\mathrm{w}|1}(N)$ approaches 0; this is in accordance with equation (17).

As far as the *majority* rule is concerned, the results established in Proposition IV.3 and in Proposition IV.4 are confirmed by the plots in Figure 3. Indeed, since for all the pairs (θ_0, θ_1) we have considered, we had $p_{\mathrm{w}|1} < 1/2$, we can see that, as expected from Proposition IV.3, the probability of wrong decision goes to 0 exponentially fast and monotonically as a function of the size of the group of the SDMs. Regarding the expected time, in all the cases, the expected decision time $\mathbb{E}[T|H_1, N, q = \lfloor N/2 \rfloor + 1]$ quickly reaches a constant value. We numerically verified that these constant values corresponded to the values predicted by the results reported in Proposition IV.4.

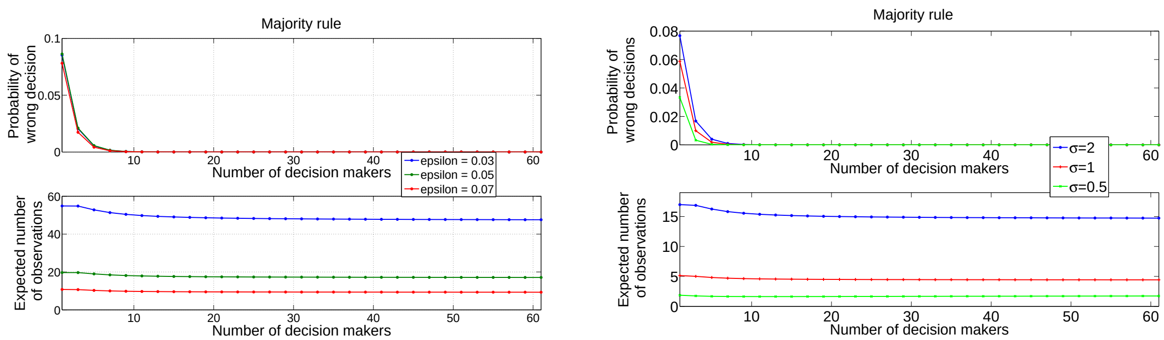

Fig. 3. Behavior of the probability of wrong decision and of the expected number of iterations required to provide a decision as the number of SDMs increases when the *majority* rule is adopted. In Figure (a) we consider the binomial distribution, in Figure (b) the Gaussian distribution.

Example V.2 (Monotonic behavior) In this example, we analyze the performance of the general q *out of* N aggregation rule, as the number of SDMs N is varied, and as the aggregation rule itself is varied. We obtained our numerical results for odd values of N ranging from 1 to 35 and for values of q comprised between 1 and $\lfloor N/2 \rfloor + 1$. Again we set the thresholds η_0 and η_1 equal to $\log(-9)$ and $\log 9$, respectively. In this example we consider only the Gaussian distribution with $\sigma = 1$. The results obtained are depicted in Figure 4, where the following monotonic behaviors appear evident:

(i) for fixed N and increasing q, both the probability of correct decision and the decision time increases;

(ii) for fixed q and increasing N, the probability of correct decision increases while the decision time decreases.

The fact that the decision time increases for fixed N and increasing q has been established in Proposition IV.6. The fact that the probability of correct decision increases for fixed N and increasing q validates the conjecture formulated at the end of Section IV-C.

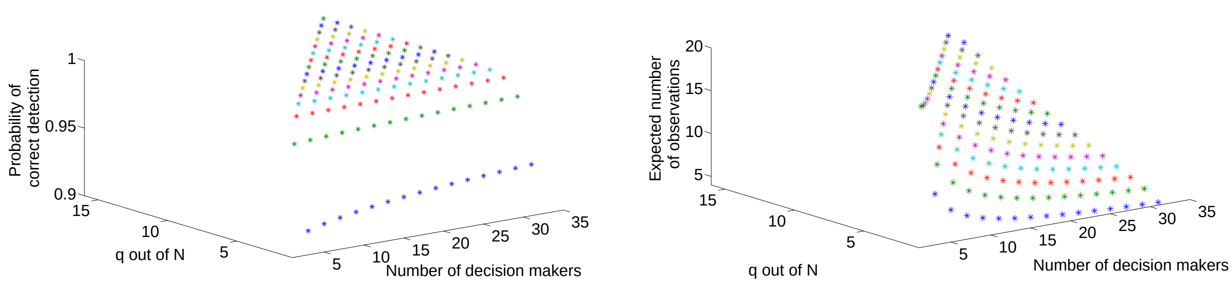

Fig. 4. Probability of correct detection (left figure) and expected decision time (right figure) for the q *out of* N rule, plotted as a function of network size N and accuracy threshold q.

Example V.3 (Fastest versus majority, at fixed group accuracy) As we noted earlier, Figures V.1-3 show that the *majority* rule increases remarkably the accuracy of the group, while the *fastest* rule decreases remarkably the expected number of iteration for the SDA to reach a decision. It is therefore reasonable to pose the following question: if the local accuracies of the SDMs were set so that the accuracy of the group is the same for both the *fastest* and the *majority* fusion rule, which of the two rules requires a smaller number of observations to give a decision. That is, at equal accuracy, which of the two rules is optimal as far as decision time is concerned.

In order to answer this question, we use a bisection on the local SDM accuracies. We apply the numerical methods presented in Proposition III.1 to find the proper local thresholds that set the accuracy of the group to the desired value $p_{\mathrm{w}|1}$. Different local accuracies are obtained for different fusion rules and this evaluation needs to be repeated for each group size N.

In these simulations, we assume the random variable X is Gaussian with variance $\sigma = 2$. The two hypotheses are $H_0 : \theta = 0$ and $H_1 : \theta = 1$. The numerical results are shown in Figure 5 and discussed below.

As is clear by the plots, the strategy that gives the fastest decision with the same accuracy varies with group size and desired accuracy. The left plot in Figure 5 illustrates that, for very high desired group accuracy, the *majority* rule is always optimal. As the accuracy requirement is relaxed, the *fastest* rule becomes optimal for small groups. Moreover, the group size at which the switch between optimal rules happens, varies for different accuracies. For example, the middle and right plot in Figure 5 illustrate that while the switch happens at $N = 5$ for a group accuracy $p^{(\mathrm{m})}_{\mathrm{w}|1} = p^{(\mathrm{f})}_{\mathrm{w}|1} = 0.05$ and at $N = 9$ for $p^{(\mathrm{m})}_{\mathrm{w}|1} = p^{(\mathrm{f})}_{\mathrm{w}|1} = 0.1$.

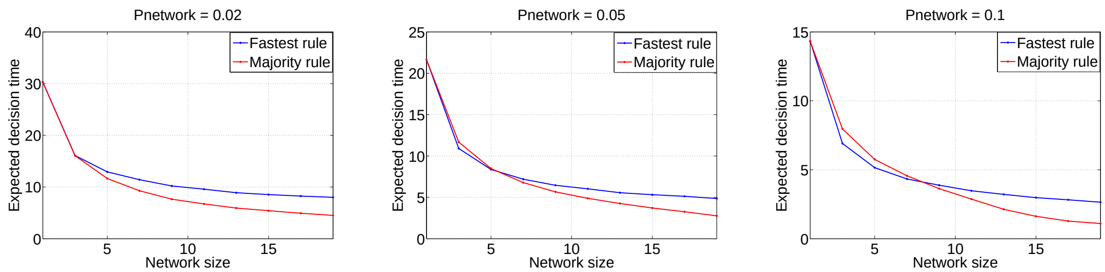

Fig. 5. Expected decision time for the *fastest* and the *majority* rules versus group size N, for various network accuracy levels.

We summarize our observations about which rule is optimal (i.e., which rule requires the least number of observations) as follows:

(i) the optimal rule varies with the desired network accuracy, at fixed network size;
(ii) the optimal rule varies with the desired network size, at fixed network accuracy; and
(iii) the change in optimality occurs at different network sizes for different accuracies.

A. Decision making in cognitive psychology revisited

In this section we point out possible relationships between our results in sequential decision aggregation (SDA) and some recent observations about mental behavior from the cognitive psychology literature. Starting with the literature review in Subsection I-D, our discussion here is based upon the following assumptions:

(i) SDA models multi-modal integration in cognitive information processing (CIP),
(ii) the number of SDMs correspond to the number of sensory modalities in CIP,
(iii) the expected decision time in the SDA setup is analogous to the reaction time in CIP, and
(iv) the decision probability in the SDA setup is analogous to the firing rate of neurons in CIP.

Under these assumptions, we point out similarities between our SDA analysis and some recent observations reported in the CIP literature. In short, the *fastest* and *majority* rules appear to emulate behaviors that are similar to the ones manifested by the brain under various conditions. We briefly mention below an example in cognitive psychology, where a parallelism might be drawn.

It is observed in cognitive information processing (CIP) that, even under the same type of stimuli, the stimuli strength affects the additivity of the neuron firing, which might end up adding up to more, less, much less, and sometimes to much more than the sum of the stimuli. These behaviors of the firing rates are called in CIP literature additive, suppressive, sub-additive or super-additive. Additionally, scientists have observed that depending on the intensity of the stimuli, various areas of the brain are activated when processing the same type of stimuli [11], [12], [13], [14], [15]. A possible explanation for these two observed behaviors is that the brain processes information in a way that maintains optimality. Indeed, our comparison in the middle and right parts of Figure 5 shows how the fastest rule is optimal when individual SDMs are highly accurate (strong and intact stimuli) and, vice versa, the majority rule is optimal when individual SDMs are relatively inaccurate (weak and degraded stimuli).

We observed in the middle and right part of Figure 5 that, for high individual accuracies, the *fastest* rule is more efficient than the *majority* rule. We reach this conclusion by noting two observations: first, smaller group sizes require higher local accuracies than larger group sizes in order to maintain the same group accuracy; second, the *fastest* rule is optimal for small groups while the *majority* rule is always optimal for larger groups. We believe that these similarities between SDA and CIP propose the possibility of explaining better, other observed behaviors in cognitive data processing if stronger links can be made between the models.

VI. Conclusion

In this work, we presented a complete analysis of how a group of SDMs 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, but a similar analysis can be made to understand other rules of interest. An important question we were able to answer, was the one relating the size of the group and the overall desired accuracy to the optimal decision rules. We were able to show that, no single rule is optimal for all group sizes or for various desired group accuracy. We are currently extending this work to cases where the individual decision makers are not identical.

References

[1] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing* (H. V. Poor and J. B. Thomas, eds.), vol. 2, pp. 297–344, 1993.

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

[3] 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.

[4] 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`.

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

[10] J. N. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Transactions on Automatic Control*, vol. 30, no. 5, pp. 440–446, 1985.

[11] F. Cook, J. Ciorciari, T. Varker, and G. J. Devilly, "Changes in long term neural connectivity following psychological trauma," *Clinical Neurophysiology*, vol. 120, no. 2, pp. 309–314, 2009.

[12] S. Werner and U. Noppeney, "Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization," *Cerebral Cortex*, vol. 11, pp. 1–14, 2009.

[13] S. Werner and U. Noppeney, "Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization," *Journal of Neuroscience*, vol. 30, no. 7, pp. 2662–2675, 2010.

[14] P. Laurienti, T. Perrault, T. Stanford, M. Wallace, and B. Stein, "On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies," *Experimental Brain Research*, vol. 166, pp. 289–297, 2005.

[15] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, "The physics of optimal decision making: A formal analysis of performance in two-alternative forced choice tasks," *Psychological Review*, vol. 113, no. 4, pp. 700–765, 2006.

[16] Y. Liu and D. Blostein, "Optimality of the sequential probability ratio test for nonstationary observations," *IEEE Transactions on Information Theory*, vol. 38, no. 1, pp. 177–182, 1992.

[17] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2538–2541, 1999.

[18] H. Okamoto, Y. Isomura, M. Takada, and T. Fukai, "Temporal integration by stochastic recurrent network dynamics with bimodal neurons," *Journal of Neurophysiology*, vol. 97, pp. 3859–3867, 2007.

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

Appendix

A. Asymptotic and monotonicity results on combinatorial sums

Some of the results provided for the *fastest* rule and for the *majority* rule are based on the following properties of the binomial expansion $(x+y)^N = \sum_{j=0}^{N} \binom{N}{j} x^j y^{N-j}$.

Lemma .1 (Properties of half binomial expansions) *For an odd number $N \in \mathbb{N}$, and for real numbers $c \in \mathbb{R}$ and $x \in \mathbb{R}$ satisfying $0 < c \leq 1$ and $0 \leq x \leq c/2$, define*

$$\underline{S}(N;c,x) = \sum_{j=0}^{\lfloor N/2 \rfloor} \binom{N}{j} x^j (c-x)^{N-j} \qquad \text{and} \qquad \overline{S}(N;c,x) = \sum_{j=\lceil N/2 \rceil}^{N} \binom{N}{j} x^j (c-x)^{N-j}.$$

The following statements hold true:

(i) *if* $0 \le x < c/2$, *then, taking limits over odd values of* N,

$$\lim_{N\to\infty} \frac{\underline{S}(N;c,x)}{c^N} = 1 \qquad \textit{and} \qquad \lim_{N\to\infty} \frac{\overline{S}(N;c,x)}{c^N} = 0;$$

(ii) *if* $x = c/2$, *then*

$$\underline{S}(N;c,x) = \overline{S}(N;c,x) = \frac{c^N}{2};$$

(iii) *if* $c = 1$ *and* $0 \le x < 1/2$, *then*

$$\overline{S}(N+2;1,x) < \overline{S}(N;1,x) \qquad \textit{and} \qquad \underline{S}(N+2;1,x) > \underline{S}(N;1,x).$$

Proof: To prove statement (i), we start with the obvious equality $c^N = (c-x+x)^N = \underline{S}(N;c,x)+\overline{S}(N;c,x)$. Therefore, it suffices to show that $\lim_{N\to\infty} \frac{\overline{S}(N;c,x)}{c^N} = 0$. Define the shorthand $h(j) := \binom{N}{j} x^j (c-x)^{N-j}$ and observe

$$\frac{h(j)}{h(j+1)} = \frac{\frac{N!}{j!(N-j)!} x^j (c-x)^{N-j}}{\frac{N!}{(j+1)!(N-j-1)!} x^{j+1} (c-x)^{N-j-1}} = \frac{j+1}{N-j}\frac{c-x}{x}.$$

It is straightforward to see that $\frac{h(j)}{h(j+1)} > 1 \iff cj - xN + c - x > 0 \iff j > \frac{xN}{c} - \frac{(c-x)}{c}$. Moreover, if $j > \frac{N}{2}$ and $0 \le x < \frac{c}{2}$, then $j - \frac{xN}{c} + \frac{c-x}{c} > \frac{N}{2} - \frac{xN}{c} + \frac{c-x}{c} \ge \frac{N}{2} - \frac{N}{2} + \frac{c-x}{c} > 0$. Here, the second inequality follows from the fact that $-\frac{xN}{c} \ge -\frac{N}{2}$ if $0 \le x < \frac{c}{2}$. In other words, if $j > \frac{N}{2}$ and $0 \le x < \frac{c}{2}$, then $\frac{h(j)}{h(j+1)} > 1$. This result implies the following chain of inequalities $h(\lceil N/2 \rceil) > h(\lceil N/2 \rceil + 1) > \cdots > h(N)$ providing the following bound on $\overline{S}(N;c,x)$

$$\frac{\overline{S}(N;c,x)}{c^N} = \frac{\sum_{j=\lceil N/2 \rceil}^{N} \binom{N}{j} x^j (c-x)^{N-j}}{c^N} < \frac{\lceil N/2 \rceil \binom{N}{\lceil N/2 \rceil} x^{\lceil N/2 \rceil} (c-x)^{\lfloor N/2 \rfloor}}{c^N}.$$

Since $\binom{N}{\lceil N/2 \rceil} < 2^N$, we can write

$$\frac{\overline{S}(N;c,x)}{c^N} < \lceil N/2 \rceil \frac{2^N x^{\lceil N/2 \rceil} (c-x)^{\lfloor N/2 \rfloor}}{c^N} = \lceil N/2 \rceil \left(\frac{2x}{c}\right)^{\lceil N/2 \rceil} \left(\frac{2(c-x)}{c}\right)^{\lfloor N/2 \rfloor}$$
$$= \lceil N/2 \rceil \left(\frac{2x}{c}\right) \left(\frac{2x}{c}\right)^{\lfloor N/2 \rfloor} \left(\frac{2(c-x)}{c}\right)^{\lfloor N/2 \rfloor}.$$

Let $\alpha = \frac{2x}{c}$ and $\beta = 2\left(\frac{c-x}{c}\right)$ and consider $\alpha \cdot \beta = \frac{4x(c-x)}{c^2}$. One can easily show that $\alpha \cdot \beta < 1$ since $4cx - 4x^2 - c^2 = -(c-2x)^2 < 0$. The proof of statement (i) is completed by noting

$$\lim_{N\to\infty} \frac{\overline{S}(N;c,x)}{c^N} < \lim_{N\to\infty} \lceil N/2 \rceil \left(\frac{2x}{c}\right) (\alpha \cdot \beta)^{\lfloor N/2 \rfloor} = 0.$$

The proof of the statement (ii) is straightforward. In fact it follows from the symmetry of the expressions when $x = \frac{c}{2}$, and from the obvious equality $\sum_{j=0}^{N} \binom{N}{j} x^j (c-x)^{N-j} = c^N$.

Regarding statement (iii), we prove here only that $\overline{S}(N+2;1,x) < \overline{S}(N;1,x)$ for $0 \le x < 1/2$. The proof of $\underline{S}(N+2;1,x) > \underline{S}(N;1,x)$ is analogous. Adopting the shorthand

$$f(N,x) := \sum_{i=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{i} x^i (1-x)^{N-i},$$

we claim that the assumption $0 < x < 1/2$ implies

$$\Delta(N,x) := f(N+2,x) - f(N,x) < 0.$$

To establish this claim, it is useful to analyze the derivative of Δ with respect to x. We compute

$$\frac{\partial f}{\partial x}(N,x) = \sum_{i=\lceil N/2 \rceil}^{N-1} i \binom{N}{i} x^{i-1} (1-x)^{N-i} - \sum_{i=\lceil N/2 \rceil}^{N-1} (N-i) \binom{N}{i} x^i (1-x)^{N-i-1} + N x^{N-1}. \tag{27}$$

The first sum $\sum_{i=\lceil N/2\rceil}^{N-1} i\binom{N}{i}x^{i-1}(1-x)^{N-i}$ in the right-hand side of (27) is equal to

$$\binom{N}{\lceil N/2\rceil}\left\lceil\frac{N}{2}\right\rceil x^{\lceil N/2\rceil-1}\left(1-x\right)^{N-\lceil N/2\rceil}+\sum_{i=\lceil N/2\rceil+1}^{N-1} i\binom{N}{i}x^{i-1}(1-x)^{N-i}.$$

Moreover, exploiting the identity $(i+1)\binom{N}{i+1}=(N-i)\binom{N}{i}$,

$$\begin{aligned}\sum_{i=\lceil N/2\rceil+1}^{N-1} i\binom{N}{i}x^{i-1}(1-x)^{N-i} &= \sum_{i=\lceil N/2\rceil}^{N-2}(i+1)\binom{N}{i+1}x^{i}(1-x)^{N-i-1}\\ &= \sum_{i=\lceil N/2\rceil}^{N-2}(N-i)\binom{N}{i}x^{i}(1-x)^{N-i-1}.\end{aligned}$$

The second sum in the right-hand side of (27) can be rewritten as

$$\sum_{i=\lceil N/2\rceil}^{N-1}(N-i)\binom{N}{i}x^{i}(1-x)^{N-i-1}=\sum_{i=\lceil N/2\rceil}^{N-2}(N-i)\binom{N}{i}x^{i}(1-x)^{N-i-1}+Nx^{N-1}.$$

Now, many terms of the two sums cancel each other out and one can easily see that

$$\frac{\partial f}{\partial x}(N,x)=\binom{N}{\lceil N/2\rceil}\lceil N/2\rceil x^{\lceil N/2\rceil-1}\left(1-x\right)^{N-\lceil N/2\rceil}=\binom{N}{\lceil N/2\rceil}\lceil N/2\rceil\left(x\left(1-x\right)\right)^{\lceil N/2\rceil-1},$$

where the last equality relies upon the identity $N-\lceil N/2\rceil=\lfloor N/2\rfloor=\lceil N/2\rceil-1$. Similarly, we have

$$\frac{\partial f}{\partial x}(N+2,x)=\binom{N+2}{\lceil N/2\rceil+1}\left(\lceil N/2\rceil+1\right)\left(x\left(1-x\right)\right)^{\lceil N/2\rceil}.$$

Hence

$$\frac{\partial \Delta}{\partial x}(N,x)=\left(x\left(1-x\right)\right)^{\lceil N/2\rceil-1}\left(\binom{N+2}{\lceil N/2\rceil+1}\left(\lceil N/2\rceil+1\right)x(1-x)-\binom{N}{\lceil N/2\rceil}\lceil N/2\rceil\right).$$

Straightforward manipulations show that

$$\binom{N+2}{\lceil N/2\rceil+1}\left(\lceil N/2\rceil+1\right)=4\frac{N+2}{N+1}\lceil N/2\rceil\binom{N}{\lceil N/2\rceil},$$

and, in turn,

$$\begin{aligned}\frac{\partial \Delta}{\partial x}(N,x)&=\binom{N}{\lceil N/2\rceil}\left\lceil\frac{N}{2}\right\rceil\left(x\left(1-x\right)\right)^{\lceil N/2\rceil-1}\left[4\frac{N+2}{N+1}x(1-x)-1\right]\\ &=: g(N,x)\left[4\frac{N+2}{N+1}x(1-x)-1\right],\end{aligned}$$

where the last equality defines the function $g(N,x)$. Observe that $x>0$ implies $g(N,x)>0$ and, otherwise, $x=0$ implies $g(N,x)=0$. Moreover, for all N, we have that $f(N,1/2)=1/2$ and $f(N,0)=0$ and in turn that $\Delta(N,1/2)=\Delta(N,0)=0$. Additionally

$$\frac{\partial \Delta}{\partial x}(N,1/2)=g(N,1/2)\left(\frac{N+2}{N+1}-1\right)>0$$

and

$$\frac{\partial \Delta}{\partial x}(N,0)=0 \qquad \text{and} \qquad \frac{\partial \Delta}{\partial x}(N,0^+)=g(N,0^+)\left(0^+-1\right)<0.$$

The roots of the polynomial $x\mapsto 4\frac{N+2}{N+1}x(1-x)-1$ are $\frac{1}{2}\left(1\pm\sqrt{\frac{1}{N+2}}\right)$, which means that the polynomial has one root inside the interval $(0,1/2)$ and one inside the interval $(1/2,1)$. Considering all these facts together, we conclude that the function $x\mapsto\Delta(N,x)$ is strictly negative in $(0,1/2)$ and hence that $f(N+2,x)-f(N,x)<0$. ■

B. Computation of the decision probabilities for a single SDM applying the SPRT test

In this appendix we discuss how to compute the probabilities

$$\left\{p_{\text{nd}|0}\right\} \cup \left\{p_{0|0}(t), p_{1|0}(t)\right\}_{t\in\mathbb{N}} \qquad \text{and} \left\{p_{\text{nd}|1}\right\} \cup \left\{p_{0|1}(t), p_{1|1}(t)\right\}_{t\in\mathbb{N}} \tag{28}$$

for a single SDM applying the classical *sequential probability ratio test* (SPRT). For a short description of the SPRT test and for the relevant notation, we refer the reader to Section V. We consider here observations drawn from both discrete and continuous distributions.

1) Discrete distributions of the Koopman-Darmois-Pitman form: This subsection review the procedure proposed in [5] for a certain class of discrete distributions. Specifically, [5] provides a recursive method to compute the exact values of the probabilities (28); the method can be applied to a broad class of discrete distributions, precisely whenever the observations are modeled as a discrete random variable of the Koopman-Darmois-Pitman form.

With the same notation as in Section V, let X be a discrete random variable of the Koopman-Darmois-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 $Z(x) = x$. Bernoulli, binomial, geometric, negative binomial and Poisson distributions are some widely used distributions of the Koopman-Darmois-Pitman form satisfying the condition $Z(x) = x$. For distributions of this form, the likelihood associated with the t-th observation $x(t)$ is given by

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

Let η_0, η_1 be the pre-assigned thresholds. Then, one can see that sampling will continue as long as

$$\frac{\eta_0 + t(A(\theta_1) - A(\theta_0))}{B(\theta_1) - B(\theta_0))} < \sum_{i=1}^{t} x(i) < \frac{\eta_1 + t(A(\theta_1) - A(\theta_0))}{B(\theta_1) - B(\theta_0))} \tag{29}$$

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 $\left\{\eta_0 + t(A(\theta_1) - A(\theta_0))\right\}/(B(\theta_1) - B(\theta_0))$ and let $\bar{\eta}_1^{(t)}$ be the largest integer smaller than $\left\{\eta_1 + t(A(\theta_1) - A(\theta_0))\right\}/(B(\theta_1) - B(\theta_0))$. Sampling will continue as long as $\bar{\eta}_0^{(t)} \leq \mathcal{X}(t) \leq \bar{\eta}_1^{(t)}$ where $\mathcal{X}(t) = \sum_{i=1}^{t} x(i)$. 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

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

and

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

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

Starting with $\mathbb{P}[\mathcal{X}(0) = 1]$, it is possible to compute recursively all the quantities $\left\{p_{i|j}(t)\right\}_{t=1}^{\infty}$ and $\mathbb{P}[\mathcal{X}(t) = \ell]$, for any $t \in \mathbb{N}$, $\ell \in [\bar{\eta}_0^{(t)}, \bar{\eta}_1^{(t)}]$, and $\left\{p_{i|j}(t)\right\}_{t=1}^{\infty}$. Moreover, if the set $\mathcal{Z}$ is finite, then the number of required computations is finite.

2) Computation of accuracy and decision time for pre-assigned thresholds η_0 and η_1: continuous distributions: In this section we assume that X is a continuous random variable with density function given by $f(x,\theta)$. As in the previous subsection, given two pre-assigned thresholds η_0 and η_1, the goal is to compute the probabilities $p_{i|j}(t) = \mathbb{P}[\text{say} H_i | H_j, T = t]$, for $i, j \in \{1, 2\}$ and $t \in \mathbb{N}$.

We start with two definitions. Let f_{λ,θ_i} and $f_{\Lambda(t),\theta_i}$ denote, respectively, the density function of the log-likelihood function λ and of the random variable $\Lambda(t)$, under the assumption that H_i is the correct hypothesis. Assume that, for a given $t \in \mathbb{N}$, the density function $f_{\Lambda(t),\theta_i}$ is known. Then we have

$$f_{\Lambda(t),\theta_i}(s) = \int_{\eta_0}^{\eta_1} f_{\lambda,\theta_i}(s-x) f_{\Lambda(t),\theta_i}(x) dx, \qquad s \in (\eta_0, \eta_1),$$

and

$$p_{i|1}(t) = \int_{\eta_0}^{\eta_1} \left(\int_{\eta_1 - x}^{\infty} f_{\lambda,\theta_i}(z) dz \right) f_{\Lambda(t),\theta_i}(x) dx, \text{ and } p_{0|i}(t) = \int_{\eta_0}^{\eta_1} \left(\int_{-\infty}^{\eta_0 - x} f_{\lambda,\theta_i}(z) dz \right) f_{\Lambda(t),\theta_i}(x) dx.$$

In what follows we propose a method to compute these quantities based on a uniform discretization of the functions λ and Λ. Interestingly, we will see how the classic SPRT algorithm can be conveniently approximated by a suitable absorbing Markov chain and how, through this approximation, the probabilities $\left\{p_{i|j}(t)\right\}_{t=1}^{\infty}$, $i, j \in \{1, 2\}$, can be efficiently computed. Next we describe our discretization approach.

First, let $\delta \in \mathbb{R}_{>0}$, $\bar{\eta}_0 = \lfloor \frac{\eta_0}{\delta} \rfloor \delta$ and $\bar{\eta}_1 = \lceil \frac{\eta_1}{\delta} \rceil \delta$. Second, for $n = \lceil \frac{\eta_1}{\delta} \rceil - \lfloor \frac{\eta_0}{\delta} \rfloor + 1$, introduce the sets

$$\mathcal{S} = \{s_1, \ldots, s_n\} \quad \text{and} \quad \Gamma = \{\gamma_{-n+2}, \gamma_{-n+3}, \ldots, \gamma_{-1}, \gamma_0, \gamma_1, \ldots, \gamma_{n-3}, \gamma_{n-2}\},$$

where $s_i = \bar{\eta}_0 + (i-1)\delta$, for $i \in \{1, \ldots, n\}$, and $\gamma_i = i\delta$, for $i \in \{-n+2, -n+3, \ldots, n-3, n-2\}$. Third, let $\bar{\lambda}$ (resp. $\bar{\Lambda}$) denote a discrete random variable (resp. a stochastic process) taking values in Γ (resp. in $\mathcal{S}$). Basically $\bar{\lambda}$ and $\bar{\Lambda}$ represent the discretization of Λ and λ, respectively. To characterize $\bar{\lambda}$, we assume that

$$\mathbb{P}\left[\bar{\lambda} = i\delta\right] = \mathbb{P}\left[i\delta - \frac{\delta}{2} \leq \lambda \leq i\delta + \frac{\delta}{2}\right], \qquad i \in \{-n+3, \ldots, n-3\},$$

and

$$\mathbb{P}\left[\bar{\lambda} = (-n+2)\delta\right] = \mathbb{P}\left[\lambda \leq (-n+2)\delta + \frac{\delta}{2}\right] \qquad \text{and} \qquad \mathbb{P}\left[\bar{\lambda} = (n-2)\delta\right] = \mathbb{P}\left[\lambda \geq (n-2)\delta - \frac{\delta}{2}\right].$$

From now on, for the sake of simplicity, we shall denote $\mathbb{P}\left[\bar{\lambda} = i\delta\right]$ by p_i. Moreover we adopt the convention that, given $s_i \in \mathcal{S}$ and $\gamma_j \in \Gamma$, we have that $s_i + \gamma_j := \bar{\eta}_0$ whenever either $i = 1$ or $i + j - 1 \leq 1$, and $s_i + \gamma_j := \bar{\eta}_1$ whenever either $i = n$ or $i + j - 1 \geq n$. In this way $s_i + \gamma_j$ is always an element of $\mathcal{S}$. Next we set $\bar{\Lambda}(t) := \sum_{h=1}^{t} \bar{\lambda}(h)$.

To describe the evolution of the stochastic process $\bar{\Lambda}$, define the row vector $\pi(t) = [\pi_1(t), \ldots, \pi_n(t)]^T \in \mathbb{R}^{1 \times n}$ whose i-th component $\pi_i(t)$ is the probability that $\bar{\Lambda}$ equals s_i at time t, that is, $\pi_i(t) = \mathbb{P}\left[\bar{\Lambda}(t) = s_i\right]$. The evolution of $\pi(t)$ is described by the absorbing Markov chain $(\mathcal{S}, A, \pi(0))$ where

- $\mathcal{S}$ is the set of states with s_1 and s_n as absorbing states;
- $A = [a_{ij}]$ is the transition matrix: a_{ij} denote the probability to move from state s_i to state s_j and satisfy, according to our previous definitions and conventions,
  - $a_{11} = a_{nn} = 1; \qquad a_{1i} = a_{nj} = 0$, for $i \in \{2, \ldots, n\}$ and $j \in \{1, \ldots, n-1\}$;
  - $a_{i1} = \sum_{s=-n+2}^{-h+1} p_s$ and $a_{in} = \sum_{s=1}^{n-2} p_s$, $h \in \{2, \ldots, n-1\}$;
  - $a_{ij} = p_{j-i} \quad i, j \in \{2, \ldots, n-1\}$;
- $\pi(0)$ is the initial condition and has the property that $\mathbb{P}[\bar{\Lambda}(0) = 0] = 1$.

In compact form we write $\pi(t) = \pi(0)A^t$.

The benefits of approximating the classic SPRT algorithm with an absorbing Markov chain $(\mathcal{S}, A, \pi(0))$ are summarized in the next Proposition. Before stating it, we provide some useful definitions. First, let $Q \in \mathbb{R}^{(n-2)\times(n-2)}$ be the matrix obtained by deleting the first and the last rows and columns of A. Observe that $I - Q$ is an invertible matrix and that its inverse $F := (I-Q)^{-1}$ is typically known in the literature as the *fundamental matrix* of the absorbing matrix A. Second let $A_{2:n-1}^{(1)}$ and $A_{2:n-1}^{(n)}$ denote, respectively, the first and the last column of the matrix A without the first and the last component, i.e., $A_{2:n-1}^{(1)} := [a_{2,1}, \ldots, a_{n-1,1}]^T$ and $A_{2:n-1}^{(n)} := [a_{2,n}, \ldots, a_{n-1,n}]^T$.

Finally, let $e_{\lfloor \frac{\eta_0}{\delta} \rfloor + 1}$ and $\mathbf{1}_{n-2}$ denote, respectively, the vector of the canonical basis of $\mathbb{R}^{n-2}$ having 1 in the $(\lfloor \frac{\eta_0}{\delta} \rfloor + 1)$-th position and the $(n-2)$-dimensional vector having all the components equal to 1 respectively.

Proposition .2 (SPRT as a Markov Chain) *Consider the classic SPRT test. Assume that we model it through the absorbing Markov chain* $(\mathcal{S}, A, \pi(0))$ *described above. Then the following statements hold:*

(i) $p_{0|j}(t) = \pi_1(t) - \pi_1(t-1)$ *and* $p_{1|j}(t) = \pi_n(t) - \pi_n(t-1)$*, for* $t \in \mathbb{N}$*;*
(ii) $\mathbb{P}[\textit{say } H_0|H_j] = e^T_{\lfloor \frac{\eta_0}{\delta} \rfloor + 1} N \bar{a}_1$ *and* $\mathbb{P}[\textit{say } H_0|H_j] = e^T_{\lfloor \frac{\eta_0}{\delta} \rfloor + 1} N \bar{a}_n$*; and*
(iii) $\mathbb{E}[T|H_j] = e^T_{\lfloor \frac{\eta_0}{\delta} \rfloor + 1} F \mathbf{1}_{n-2}$.