

# Distribution Inequalities for Parallel Models of Reaction Time with an Application to Auditory Profile Analysis

Hans Colonius

*Universität Oldenburg, Oldenburg, Germany*

and

Wolfgang Ellermeier

*Universität Regensburg, Regensburg, Germany*

---

Inequalities on reaction time distribution functions for parallel models with an unlimited capacity assumption are presented, extending previous work on first-terminating and exhaustive stopping rules to second-terminating processes. This extension thus generates transitions between first-terminating and exhaustive models that might be of interest in situations in which observers behave as if collecting more evidence before a decision is made. Moreover, a generalization of the inequalities is derived and tested in an auditory profile analysis task in which subjects have to decide whether two multitone complexes are of the same or of different spectral shape. © 1997 Academic Press

---

## 1. INTRODUCTION

Models of information processing in a perceptual/cognitive task commonly assume that reaction time (RT) can be decomposed into several subprocesses. It has been shown by a number of authors that only if additional assumptions are introduced, is it possible to infer from RT measurements whether these subprocesses operate in parallel or serially (Townsend, 1972, 1974; Vorberg, 1977; Vorberg & Ulrich, 1987). In addition to processing architecture, a problem of considerable importance has been the stopping rule the information processing system employs when sufficient information has been acquired to make a correct response. Many of these results have been reviewed in Townsend and Ashby (1983), Townsend (1990), and Luce (1986). The aim of the present paper is (1) to extend results of Colonius and Vorberg (1994) on stopping rules for parallel models in various ways, and (2) to present an application of these results to an auditory discrimination task. For recent results regarding stopping rules for serial models, see Townsend and Colonius (in press).

Send correspondence and reprint requests to Hans Colonius, Institut für Kognitionsforschung, Universität Oldenburg, FB5-A6, D-26111 Oldenburg/Germany. E-mail: colonius@psychologie.uni-oldenburg.de. Telephone/fax: +49 441 798 5158.

We describe a stochastic model of a parallel system with  $n$  processing channels by a collection of nonnegative random variables  $T_1, T_2, \dots, T_n$  which refer to the processing durations of the channels. The term *channel* is used here in a broad sense. For example, in a redundant-signals task, where human observers monitor two or more sources of information for a target stimulus, the  $T_k$  refer to the signal-specific processing times. In the context of the auditory discrimination paradigm considered below, channel processing time will refer to the time to measure the level of a signal component within a multitone complex.

Channels may interact with each other; thus, processing times  $T_k, k = 1, \dots, n$ , are allowed to be stochastically dependent. In accordance with the terminology proposed by Colonius and Vorberg (1994), a parallel system is called *first-terminating* if it initiates a response as soon as the first channel finishes processing. Thus, the contribution of the channels' processing time to overall reaction time is the time needed by the fastest channel to finish processing, i.e., the minimum of the  $T_k$ . Alternatively, a parallel system is called *exhaustive* if it cannot initiate a response unless all its channels have finished. In this case, the contribution to overall reaction time is given by the time of the slowest channel, i.e., the maximum of the  $T_k$ . In general, the overall reaction time may comprise components like stimulus encoding time, response selection time, and motor execution time; all of these additional components are often summarized by adding another random variable to the maximum (or minimum) random variable.

While parallel models with a first-terminating or an exhaustive stopping rule are plausible mechanisms in many situations, note that these rules only constitute two extreme cases of a more general situation where overall reaction time is determined by  $k$  channels having finished ( $k = 1, \dots, n$ ). Depending on the specific conditions of the experiment or on instructions given to subjects, the general  $k$ th-terminating

rule ( $k$  different from 1 or  $n$ ) may yield a more realistic description of the data.

In the next section, after introducing some necessary terminology, relevant results from Colonius and Vorberg (1994) are summarized. Their distribution inequalities in the first-terminating case are then extended to the second terminating ( $k = 2$ ) case. A subsequent section presents an application of these ideas to the auditory profile discrimination task. Finally, predictions of the models are compared to a preliminary set of data.

## 2. DISTRIBUTION INEQUALITIES: FIRST-TERMINATING AND EXHAUSTIVE MODELS

We start with

**DEFINITION 1.** Let  $T_1, T_2, \dots, T_n$  be jointly distributed random variables; the corresponding *order statistics* are the  $T_i$ 's arranged in nondecreasing order

$$T_{1:n} \leq T_{2:n} \leq \dots \leq T_{n:n}$$

Specifically,  $T_{k:n}$  is called the *kth-order statistic*. When actual equalities apply, we do not make any requirement about which variable should precede the other one.

This definition does not require that the  $T_i$ 's be identically distributed, or that they are independent, or that densities exist. In the reaction-time modeling context, though, existence of a multivariate density for the  $T_1, T_2, \dots, T_n$  can usually be taken for granted. Many classical results dealing with order statistics were originally derived in the i.i.d. case (*independent identically distributed*) with common continuous (cumulative) distribution (see Arnold, Balakrishnan, & Nagaraja, 1992; Galambos, 1978).

**DEFINITION 2.** Let  $T_1, T_2, \dots, T_n$  denote the random channel processing times of a parallel reaction time model. The model is called *parallel kth-terminating* if its reaction time (RT) is equal to the  $k$ th-order statistic of  $T_1, T_2, \dots, T_n$ , i.e.,

$$\text{RT} = T_{k:n}.$$

In particular, the model is called *parallel first-terminating* if

$$\text{RT} = \min_{i=1, \dots, n} T_i = T_{1:n}.$$

It is called *parallel exhaustive* if

$$\text{RT} = \max_{i=1, \dots, n} T_i = T_{n:n}.$$

As in Colonius and Vorberg (1994), we confine attention to parallel models with unlimited processing capacity. By this we mean that the system allots the same amount of

capacity to a given channel no matter how many additional channels operate at the same time. This notion of unlimited capacity can be made precise by considering the joint probability distribution functions of the times  $T_i$  over experimental conditions with different sets of active channels. Suppose we have a subset of  $m$  out of the  $n$  channels of the system. Unlimited capacity stipulates that their (marginal) distributions are the same no matter how many and which of the remaining  $n - m$  channels are active.

**DEFINITION 3.** Let  $B = \{i_1, i_2, \dots, i_m\}$  ( $m \leq n$ ) be the set of channels active under a given experimental condition and let  $P_B$  denote the corresponding probability measure implied by the  $m$ -channel model; and  $n$ -channel parallel model is said to have *unlimited capacity* if

$$\begin{aligned} P_B(T_{i_1} \leq t_1, T_{i_2} \leq t_2, \dots, T_{i_m} \leq t_m) \\ = P(T_{i_1} \leq t_1, T_{i_2} \leq t_2, \dots, T_{i_m} \leq t_m) \end{aligned}$$

for all  $B \subset \{1, 2, \dots, n\}$  and all  $t_1, t_2, \dots, t_m \in \mathfrak{R}^+$ , where  $P$  denotes the probability measure corresponding to the model where all  $n$  channels are active.

Note that this definition of unlimited capacity neither implies nor is implied by stochastic independence of the processing times. An analogous definition was proposed by Vorberg (1981) for serial models. In many experimental situations, the unlimited-capacity assumption can be expected to hold at most for a limited number of parallel channels. It seems plausible in systems involving highly automatized processing like short-term memory scanning or, as discussed later, auditory profile-analysis tasks.

In a parallel first-terminating system with  $n$  channels a response is initiated as soon as any of the channels finishes processing. Intuitively, adding another channel to the race cannot but increase—or, at least, keep constant—the probability that the first channel finishes before some time  $t$  if unlimited capacity is assumed to hold because it is the additional channel which may be the fastest. This intuition is corroborated by the first inequality (lower bound) in the theorem below taken from Colonius and Vorberg (1994). On the other hand, the probability of finishing processing with  $n$  channels before time  $t$  is bounded from above by a linear combination of the corresponding probabilities with  $n - 1$  and  $n - 2$  channels operating. This bound is specified by the right-hand side of the first inequality in the theorem to follow. Analogously, adding another channel in a parallel exhaustive system cannot but decrease the probability that processing will be finished before some time  $t$ , since this added channel may take longer to finish than any of the others. Again, however, this probability is bounded from below as specified in the second inequality in the theorem.

Consistent with the order statistics notation, distribution functions for first-terminating and exhaustive parallel

models with  $n$  channels will be denoted by  $F_{1:n}$  and  $F_{n:n}$ , respectively; thus,

$$F_{1:n}(t) = P(T_{1:n} \leq t) \quad (\text{first-terminating } n\text{-channel process})$$

and

$$F_{n:n}(t) = P(T_{n:n} \leq t) \quad (\text{exhaustive } n\text{-channel process}).$$

Moreover, we write  $F_{1:n-1}^{(i)}(t)$ ,  $F_{1:n-2}^{(ij)}(t)$ , etc. for the distribution function if all channels but channel  $i$ , all channels but channels  $i, j$ , etc. are active, respectively. Thus, using the notation of Definition 3 with  $C = \{1, 2, \dots, n\}$  we have, for example,

$$\begin{aligned} F_{n-1:n-1}^{(i)}(t) \\ = P_{C \setminus \{i\}}(T_1 \leq t, T_2 \leq t, \dots, T_{i-1} \leq t, T_{i+1} \leq t, \dots, T_n \leq t). \end{aligned}$$

By the unlimited capacity assumption, this equals

$$P(T_1 \leq t, T_2 \leq t, \dots, T_{i-1} \leq t, T_{i+1} \leq t, \dots, T_n \leq t).$$

**THEOREM 1** (Colonius & Vorberg, 1994). *For an  $n$ -channel parallel model ( $n > 2$ ) with unlimited capacity with  $1 \leq i, j \leq n$ ,  $i \neq j$ , the following holds for all real  $t (t > 0)$ :*

$$\begin{aligned} \max_i F_{1:n-1}^{(i)}(t) \\ \leq F_{1:n}(t) \leq \min_{i,j} [F_{1:n-1}^{(i)}(t) + F_{1:n-1}^{(j)}(t) - F_{1:n-2}^{(ij)}(t)] \quad (1) \end{aligned}$$

(first-terminating model) and

$$\begin{aligned} \max_{i,j} [F_{n-1:n-1}^{(i)}(t) + F_{n-1:n-1}^{(j)}(t) - F_{n-2:n-2}^{(ij)}(t)] \\ \leq F_{n:n}(t) \leq \min_i F_{n-1:n-1}^{(i)}(t) \quad (2) \end{aligned}$$

(exhaustive model).

**DEFINITION 4.** If all marginal distributions  $F_{1:n-1}^{(i)}$ ,  $F_{1:n-2}^{(ij)}$ ,  $F_{n-1:n-1}^{(i)}$ ,  $F_{n-2:n-2}^{(ij)}$ , etc. are independent of the choice of  $i$  and  $j$ , the parallel system is called *marginally invariant*.

For marginally invariant parallel systems, the inequalities take a specially simple form.

**COROLLARY 1.** *For a marginally invariant parallel system, the inequalities of Theorem 1 become*

$$F_{1:n-1}(t) \leq F_{1:n}(t) \leq 2F_{1:n-1}(t) - F_{1:n-2}(t) \quad (3)$$

and

$$2F_{n-1:n-1}(t) - F_{n-2:n-2}(t) \leq F_{n:n}(t) \leq F_{n-1:n-1}(t). \quad (4)$$

Although not much is known about the sharpness of the upper and the lower bounds in the above inequalities, in general, it should not escape the reader that these inequalities are valid without any specific assumptions concerning the stochastic independence of the channel processing times or the parametric form of their distribution. In a recent study, Ashby, Tein, and Balakrishnan (1993) employed the above result to rule out a large class of unlimited-capacity, parallel exhaustive models for a memory-scanning task.

### 3. DISTRIBUTION INEQUALITIES FOR SECOND-TERMINATING PARALLEL MODELS

Several experimental situations are conceivable where response times, rather than being controlled by the first channel terminating its processing, are in fact determined by the second, by the third, or by some  $k$ th channel's processing time. Consider, for example, a same-different matching paradigm, where the subject has to decide as quickly as possibly whether two multicomponent stimuli are identical or not. If the subject is encouraged to avoid incorrect responses, a "different" response may not be given before at least  $k$ ,  $k \geq 2$ , different components have been detected. A more elaborate example, an auditory discrimination paradigm (profile analysis), will be presented in a later section.

We now investigate the second-terminating case ( $k = 2$ ). To derive inequalities analogous to those above, some general probability inequalities have to be established. These are different from those called upon in proving Theorem 1 (see Colonius & Vorberg, 1994) and may possibly be utilized to extend our results to the general case (i.e.,  $k \neq 1, 2$ , or  $n$ ) which will not be treated here. A reader not interested in the proofs may skip the following subsection without loss of continuity.

#### 3.1. An Improved Bonferroni Inequality

The classical *Bonferroni inequalities* were introduced by Bonferroni (1937). Let  $A_1, A_2, \dots, A_n$  be a sequence of events in a probability space. For example, the second-order Bonferroni inequality is

$$\begin{aligned} \sum_{i=1}^n P(A_i) - \sum_{i>j} \sum_{j=1}^n P(A_i \cap A_j) \\ \leq P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i). \quad (5) \end{aligned}$$

These bounds can be quite inaccurate and, therefore, improved bounds, referred to as *Bonferroni-type* inequalities,

have been developed (see, e. g., Glaz, 1992; Worsley, 1982). The basic idea is to express

$$A = A_1 \cup (A_2 \cap A_1^c) \cup (A_3 \cap A_2^c \cap A_1^c) \cup \dots \cup (A_n \cap A_{n-1}^c \cap \dots \cap A_1^c). \quad (6)$$

This immediately implies the following Bonferroni-type upper bound that will be used in the next section:

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i) - \sum_{i=1}^{n-1} P(A_i \cap A_{i+1}). \quad (7)$$

### 3.2. An Inequality for Second-Terminating Parallel Models

The response time distribution of a second-terminating parallel model with  $n$  channel processing times  $T_1, T_2, \dots, T_n$  is defined by the second-order statistic; i.e.,

$$\begin{aligned} P_n(T_{2:n} \leq t) &= F_{2:n}(t) \\ &= P_n \quad (\text{at least two of } T_1, T_2, \dots, T_n \\ &\quad \text{are less than or equal to } t), \end{aligned}$$

where  $P_n$  refers to the probability measure corresponding to the joint distribution of  $T_1, T_2, \dots, T_n$ . The goal here is to find upper and lower bounds for this distribution function by utilizing the corresponding distributions for models with  $n-1$  and  $n-2$  channels under the assumption of unlimited capacity. Let us write

$$\begin{aligned} P_{n-1}(T_{2:n-1} \leq t) &= F_{2:n-1}^{(t)} \\ &= P_{n-1} \quad (\text{at least two of } T_1, T_2, \dots, T_{i-1}, \\ &\quad T_{i+1}, \dots, T_n \text{ are less than or} \\ &\quad \text{equal to } t). \end{aligned}$$

By unlimited capacity, the subscripts  $n, n-1$ , etc. can be dropped from the probability measure  $P$ . Let

$$B_{kl}(t) = \{T_k \leq t \cap T_l \leq t\}, \quad A(t) = \bigcup_{\substack{k < l \\ k, l = 1, \dots, n}} B_{kl}(t).$$

For notational convenience, we drop the variable  $t$  in the set notation. Then, obviously,

$$F_{2:n}(t) = P(A).$$

Moreover, writing

$$A^{(i)} = \bigcup_{\substack{k < l \\ k, l = 1, \dots, n \\ k, l \neq i}} B_{kl}$$

yields

$$F_{2:n-1}^{(i)}(t) = P(A^{(i)}).$$

The following technical lemma will allow us to utilize the Bonferroni-type inequality referred to in the last subsection.

LEMMA 1. *With the notation introduced above, for any pairwise different  $p, q, r$  from  $\{1, \dots, n\}$*

$$A = A^{(p)} \cup A^{(q)} \cup A^{(r)}.$$

*Proof.* Consider the set  $B_{ij}$  for arbitrary  $i, j$  ( $i \neq j$ ) from  $\{1, \dots, n\}$ . If  $i, j \neq p, q, r$ , then by definition  $B_{ij}$  is contained in  $A^{(p)} \cap A^{(q)} \cap A^{(r)}$ . If exactly one of  $i, j$  is equal to either  $p, q$ , or  $r$ , then  $B_{ij}$  is contained in the intersection of two of the sets  $A^{(p)}, A^{(q)}, A^{(r)}$ . For example, if  $i=p$ , then  $B_{ij} \subset A^{(q)} \cap A^{(r)}$ . Finally, if both  $i$  and  $j$  are equal to two of  $p, q, r$ , then  $B_{ij}$  is contained in exactly one of the sets  $A^{(p)}, A^{(q)}, A^{(r)}$ . For example, let  $i=p$  and  $j=q$ . Then,  $B_{ij} \subset A^{(r)}$ . Q.E.D.

We now consider the case where two of the channels,  $i$  and  $j$ , say, are dropped. Extending the above notation, let us write

$$A^{(ij)} = \bigcup_{\substack{k < l \\ k, l = 1, \dots, n \\ k, l \neq i, j}} B_{kl}.$$

LEMMA 2. *With the notation introduced above, for any  $p, q$  ( $p \neq q$ ) from  $\{1, \dots, n\}$*

$$A^{(pq)} \subset A^{(p)} \cap A^{(q)}.$$

*Proof.* By construction,  $A^{(pq)} \subset A^{(p)}$  and  $A^{(pq)} \subset A^{(q)}$ . Q.E.D.

We are now ready to utilize the Bonferroni-type inequality (7) to derive an upper bound for the RT distribution function with  $n$  channels,  $F_{2:n}(t)$ . By Lemma 1 and (7),

$$\begin{aligned} P(A) &= P(A^{(p)} \cup A^{(q)} \cup A^{(r)}) \\ &\leq P(A^{(p)}) + P(A^{(q)}) + P(A^{(r)}) \\ &\quad - P(A^{(p)} \cap A^{(q)}) - P(A^{(q)} \cap A^{(r)}) \\ &\leq P(A^{(p)}) + P(A^{(q)}) + P(A^{(r)}) \\ &\quad - P(A^{(pq)}) - P(A^{(qr)}) \end{aligned} \quad (8)$$

with the last inequality following from Lemma 2. Translating the above results from the set notation into probability distributions leads to the desired upper and lower bounds for  $F_{2:n}(t)$ .

**THEOREM 2.** *For an  $n$ -channel parallel model ( $n > 2$ ) with unlimited capacity, the following holds:*

$$F_{2:n-1}^{(p)}(t) \leq F_{2:n}(t) \leq F_{2:n-1}^{(p)}(t) + F_{2:n-1}^{(q)}(t) + F_{2:n-1}^{(r)}(t) - F_{2:n-2}^{(pq)}(t) - F_{2:n-2}^{(qr)}(t). \quad (9)$$

Obviously, the upper (lower) bound can be sharpened by minimizing (maximizing) over the indices  $p, q, r$ .

**COROLLARY 2.** *Under the conditions of Theorem 1, for all  $t (t > 0)$  with pairwise different  $p, q, r, 1 \leq p, q, r \leq n$ ,*

$$\max_p F_{2:n-1}^{(p)}(t) \leq F_{2:n}(t) \leq \min_{(p, q, r)} \{F_{2:n-1}^{(p)}(t) + F_{2:n-1}^{(q)}(t) + F_{2:n-1}^{(r)}(t) - F_{2:n-2}^{(pq)}(t) - F_{2:n-2}^{(qr)}(t)\}, \quad (10)$$

where the minimum is taken over all triples  $(p, q, r)$ .

Under marginal invariance (Definition 4), the inequalities take a specially simple form.

**COROLLARY 3.** *Under marginal invariance, the inequalities become*

$$F_{2:n-1}(t) \leq F_{2:n}(t) \leq 3F_{2:n-1}(t) - 2F_{2:n-2}(t). \quad (11)$$

### 3.3. First-Terminating versus Second-Terminating Parallel Models

It is instructive to compare the response time distributions under the first and second-terminating stopping rule. For simplicity, we restrict this discussion to the marginally invariant case. Obviously, for any number  $m$  of channels  $F_{2:m}(t) \leq F_{1:m}(t)$ . Combining this with the lower bounds from Theorem 1 and Corollary 2 yields, for example,

$$F_{2:n-1}(t) \leq \min\{F_{1:n-1}(t), F_{2:n}(t)\} \leq \max\{F_{1:n-1}(t), F_{2:n}(t)\} \leq F_{1:n}(t). \quad (12)$$

From an empirical point of view, inequalities like (12) will only be relevant if there is some experimental manipulation guaranteeing—or, at least, suggesting—a switch from first-terminating to second-terminating processing or vice versa. A more common situation will be one with  $n$  parallel channels operating where no information about the stopping rule is available a priori. In other words, a response time distribution  $G_n(t)$  is observable (rather, estimable), possibly for different numbers of channels  $n$ . In that case, the only testable condition derivable from (12) is  $G_{n-1}(t) \leq G_n(t)$ , which is identical to the lower bound in Corollary 1. While Corollary 1 can be utilized to discriminate between first-terminating and exhaustive processing, Corollary 3 is potentially useful to distinguish between first- and second-terminating model. To see this, consider the upper bound of

Corollary 1 valid for the first-terminating stopping rule. In terms of the observable RT distribution, this will be

$$G_n(t) \leq 2G_{n-1}(t) - G_{n-2}(t). \quad (13)$$

On the other hand, the upper bound of Corollary 3 valid for the second-terminating stopping rule is

$$G_n(t) \leq 3G_{n-1}(t) - 2G_{n-2}(t) = 2G_{n-1}(t) - G_{n-2}(t) + \{G_{n-1}(t) - G_{n-2}(t)\}. \quad (14)$$

For both first- and second-terminating models,  $\{G_{n-1}(t) - G_{n-2}(t)\}$  will be nonnegative. Thus, the upper bound in (13) is stricter than the upper bound in (14). Even if data are not consistent with the first-terminating stopping rule, they may still be consistent with the second-terminating model.

## 4. APPLICATION: DISCRIMINATION OF AUDITORY PROFILES

This section presents an application of the distribution inequalities approach to an auditory model. It is an elaborated version of the first-terminating and exhaustive case, respectively. The preliminary data are only meant to illustrate the approach. A more comprehensive presentation of the experiment is deferred to another paper.

With the exception of a few applications to loudness and intensity discrimination (e.g., Kohfeld *et al.*, 1981; Burbeck & Luce, 1982), there are hardly any studies in psychoacoustics that have used reaction time in order to reveal mechanisms of auditory processing. However, several paradigms in hearing research lend themselves to such analysis, and the one we chose to investigate is *auditory profile analysis* (Green, 1988, 1992). Here the profile discrimination paradigm serves to illustrate the applicability of our theoretical results and also extends them in a non-trivial way. In particular, we examine and modify a model originally proposed for search in semantic memory by Vorberg and Schmidt (see Vorberg, Colonius, & Schmidt, 1989).

### 4.1. Modeling Auditory Profile Discrimination

In a profile analysis task, the subject is to discriminate broadband sounds on the basis of spectral shape, typically by discriminating a flat multitone spectrum from one in which one or several components have been incremented relative to the others. It is assumed that the subject accomplishes this task by first performing a critical-band analysis and then comparing levels across several auditory filters. The temporal dynamics of such analysis (whether processing is serial, parallel, capacity-limited, etc.) have not been addressed in models of spectral-shape discrimination, but

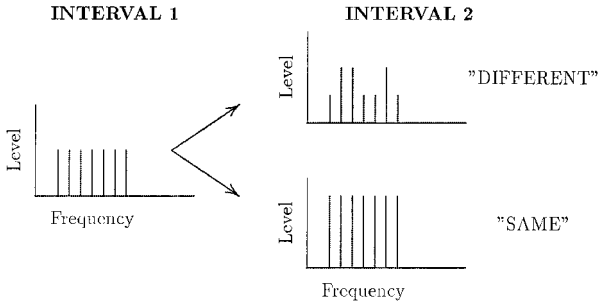


FIG. 1. Trial sequence in the same-different profile discrimination experiment (see text).

there is some evidence that processing times may play a role. For example, Dai and Green (1993) varied the duration of multitone complexes from 5 ms to 1 s and found that profile discrimination thresholds for a 21-component complex suffered more as duration was decreased than the corresponding thresholds for a 3-component complex. While they propose a somewhat different explanation, this outcome is consistent with the idea that given a short time-window and a great number of stimulus components not all of them can be sufficiently analyzed.

In order to model profile discrimination using reaction times as data, we make the following assumptions: (1) In order to detect a profile change, the auditory system performs relative level comparisons across the spectrum. (2) The time it takes to obtain level “measurements” from individual frequency channels can be considered a random variable. (3) Level is assessed in parallel in all frequency channels excited.

With  $m$  (incremented) signal components and  $n$  (nonincremented) background components present in a given auditory profile, all it would require to discriminate a “flat” from an uneven profile (see Fig. 1) would be to measure the level of at least one signal component and of at least one background component.<sup>1</sup> This is an extension of the first-terminating model considered above. To be specific, let  $S_1, \dots, S_m$  and  $B_1, \dots, B_n$  denote the random times needed to make the level “measurements” of the  $m$  signal components and the  $n$  background components, respectively. Overall reaction time for this  $(m, n)$ -channel parallel model could then be determined by the fastest measurement obtained for the signal components and by the fastest measurement for the background components, respectively, and the completion of both; i.e.,

$$\begin{aligned} RT_{m,n}^{[1,1]} &= \max\{\min\{S_1, \dots, S_m\}, \min\{B_1, \dots, B_n\}\} \\ &= \max\{S_{1:m}, B_{1:n}\}, \end{aligned} \quad (15)$$

<sup>1</sup> Since in a typical profile experiment, the overall intensity level of the multitone complex is roved over a certain dB range, it is not sufficient for the auditory mechanism to measure just one signal component (see Fig. 1).

where the superscript in  $RT_{m,n}^{[1,1]}$  refers to the  $(k, l) = (1, 1)$  (first-terminating) stopping rule applied to both the signal and the background components. It is equally conceivable, however, that the auditory system collects evidence from all frequency channels, before a decision is made, that is, the *exhaustive*  $(m, n)$ -channel model, which may be written as

$$\begin{aligned} RT_{m,n}^{[m,n]} &= \max\{S_1, \dots, S_m, B_1, \dots, B_n\} \\ &= \max\{S_{m:m}, B_{n:n}\}. \end{aligned} \quad (16)$$

Extending the notation from the previous sections, the distribution function for  $RT_{m,n}^{[1,1]}$  and  $RT_{m,n}^{[m,n]}$  will be written  $F_{m,n}^{[1,1]}(t)$  and  $F_{m,n}^{[m,n]}(t)$ , respectively, etc. To keep matters simple, marginal invariance will be assumed separately, with respect to the signal components and the background components. Extension to the general case is straightforward. The following extends Corollary 1 to the  $(m, n)$ -channel model.

**COROLLARY 4.** *For an  $(n, m)$ -channel parallel model ( $n > 2, m > 2$ ) with unlimited capacity and under marginal invariance, the following holds for all  $t(t > 0)$ ,*

$$\begin{aligned} &\max\{F_{m-1,n}^{[1,1]}(t), F_{m,n-1}^{[1,1]}(t)\} \\ &\leq F_{m,n}^{[1,1]}(t) \leq \min\left\{ \begin{aligned} &2 * F_{m-1,n}^{[1,1]}(t) - F_{m-2,n}^{[1,1]}(t), \\ &2 * F_{m,n-1}^{[1,1]}(t) - F_{m,n-2}^{[1,1]}(t) \end{aligned} \right\} \end{aligned} \quad (17)$$

(*first-terminating processes*) and

$$\begin{aligned} &F_{m-1,n}^{[m-1,n]}(t) + F_{m,n-1}^{[m,n-1]}(t) - F_{m-1,n-1}^{[m-1,n-1]}(t) \\ &\leq F_{m,n}^{[m,n]}(t) \leq \min\{F_{m-1,n}^{[m-1,n]}(t), F_{m,n-1}^{[m,n-1]}(t)\} \end{aligned} \quad (18)$$

(*exhaustive processes*).

*Proof* (Sketch). For the first-terminating part, define

$$S_i^* = \max\{S_i, B_{1:n}\}, \quad B_j^* = \max\{S_{1:m}, B_j\}.$$

It is not hard to see that

$$S_{1:m}^* = B_{1:n}^* = RT_{m,n}^{[1,1]}.$$

Applying the first-terminating part of Corollary 1 to both  $S_{1:m}^*$  and  $B_{1:n}^*$  and minimizing over the two upper bounds (maximizing over the two lower bounds) yields the result for the first-terminating part. The exhaustive part is trivial. Q.E.D.

Note that these inequalities call for six experimental conditions to be investigated, in which one or two of each set of

components (signal or background) are taken out, in order to compute upper and lower reaction-time boundaries for the baseline conditions  $F_{m,n}^{[1,1]}$  or  $F_{m,n}^{[m,n]}$ .

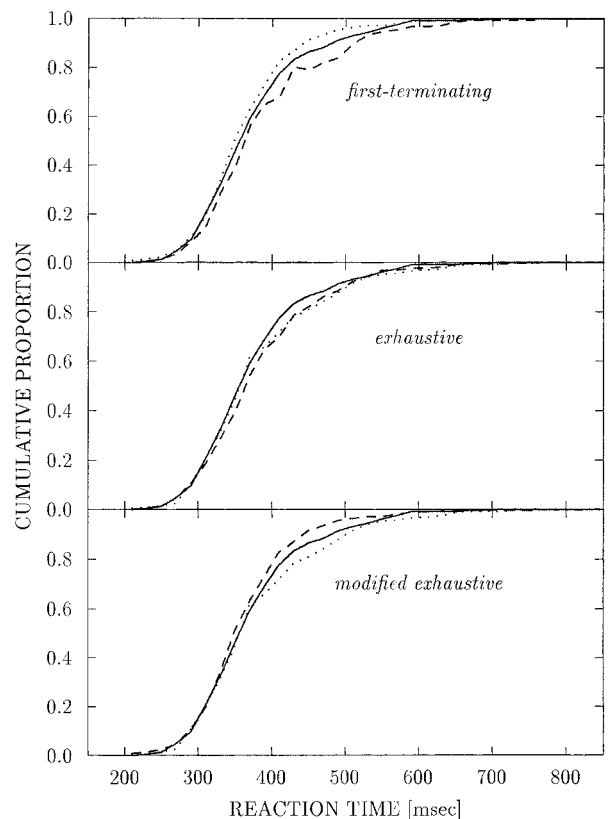
#### 4.2. Profile Discrimination Experiment

For details of the experimental setup and the full results, the reader is referred to Ellermeier and Colonius (1995, in preparation). In order to obtain reaction-time distributions for the conditions in the inequalities of Corollary 4, multi-tone complexes consisting of up to eight sinusoids equally spaced in log frequency in the range from 267 to 2400 Hz, for which detectability is known to change very little (Bernstein & Green, 1988), were generated. To generate 7-component or 6-component stimuli, the highest or the two outer components of the 8-component stimulus so defined were omitted. Which of the stimulus components carried one of the  $m$  increments, was randomly determined on a trial-by-trial basis. The starting phase of each component was randomly selected for each presentation in order to render the temporal waveform uninformative. All stimuli had a duration of 200 ms, including linear 10 ms rise and decay ramps. The mean level was set at 50-dB SPL per component. To prevent subjects from using absolute level as a cue to the presence of a signal, the overall level was roved over a 20-dB range from presentation to presentation.

*Procedure.* As depicted in Fig. 1, the trial sequence started with the presentation of the (flat) standard profile. Following an exponentially distributed foreperiod which started 300 ms after stimulus offset, had an expected value of 500 ms, and was truncated 2 s later, the actual test stimulus was presented. It either had the same flat profile, or one of the six profile alterations defined in (17) and (18), with  $m = n = 4$ . Thus an  $(m - 2, n)$  trial, for example, consisted of presentation of a 6-component standard followed by a test stimulus, in which two components were incremented relative to the remaining four. The magnitude of the increment was fixed, producing a level difference  $\Delta L$  of 10 dB between signal and background components in all conditions investigated. On each trial, the subject had to decide as quickly as possible, if the two sounds had the “same” (flat–flat) or a “different” (flat–altered) spectral profile by pressing one of two response buttons. Reaction time was measured from the onset of the test stimulus. After considerable training, about 3000 trials were collected per subject, distributed over several sessions, thus yielding roughly 500 trials in each of the six conditions.

*Results.* The critical tests implied by the inequalities shall be illustrated using the data of one representative subject. First, cumulative reaction-time distribution functions with bin widths of 10 ms were generated from the correct trials in all stimulus configurations. Subsequently, the outcome for condition  $(n, m)$  was plotted along with the upper

and lower boundaries computed from the other conditions. To evaluate the first-terminating model, for example, the lower boundary is computed from the two experimental conditions which involve taking out one component only (signal or background), as prescribed in the left side of (17). The required maximum curve is determined by scanning through all of the RT bins individually, determining which of the two cumulative distribution functions produces the higher value and taking that value as one of the points defining the lower boundary. The upper boundary for the first-terminating model is computed in an analogous manner from the right side of (17) from conditions involving the omission of one or two sinusoidal components, respectively. As is illustrated in Fig. 2 (top and middle graphs), both the first-terminating and the exhaustive model as specified in Corollary 4 failed to predict the data with more severe violations being evident in the first-terminating case. Furthermore, applying a second-terminating stopping rule as in Corollary 3 did not improve the situation. Therefore, the



**FIG. 2.** Cumulative RT distribution function obtained from a single subject in the profile analysis task (solid line) plotted along with the upper (dashed) and lower (dotted) bounds computed according to three parallel-processing models. The top panel shows RT boundaries for the “first-terminating” model (Inequality (17)), the middle panel for the “exhaustive” model (Inequality (18)), and the bottom panel for the “modified exhaustive” model discussed in the text. Note that only for the modified exhaustive model do the data fall between the boundaries thus specified.

exhaustive model was modified by dropping the minimum requirement in the upper bound. Instead, the condition involving  $m - 1$  signal components and  $n$  background components alone was used to determine the upper bound for the critical distribution function. The resulting distributions, shown in Fig. 2 (bottom graph), suggest the modified exhaustive model to give a good description for the subject.

Dropping any of the minimum or maximum requirements in the inequalities implies stating that omitting a signal component cannot be treated in the same way as omitting a nonsignal component. This is not all that surprising and may be related to recent work from our laboratory showing that decrements in a spectral profile are less detectable than increments of equal magnitude (Ellermeier, 1996).

Thus while the present data seem to support a model according to which all stimulus components have to be analyzed before a decision on spectral shape can be made, it should be noted that the effects favoring this model are rather small and still await significance testing. It is encouraging to see, however, that a number of models can be rejected just on the basis of reversed upper and lower bounds, for example. Further work might explore other detection and discrimination paradigms to find instances in which the more sophisticated first- or second-terminating models hold.

## 5. CONCLUDING REMARKS

In this paper, distribution inequalities for parallel first-terminating and exhaustive models presented in Colonius and Vorberg (1994) have been extended to the second-terminating case. Moreover, the usefulness of these inequalities has been illustrated by a theoretical and an empirical analysis of an auditory profile discrimination task. However, a large number of issues, both theoretical and practical, remain to be solved. Let us mention just some of them.

First, results on the general  $k$ -terminating case could be developed, along the lines of the arguments given here for the  $k = 2$  case. Second, in the context of the  $(m, n)$ -channel model, developed here for the auditory profile discrimination task, derivations for the second-terminating case are straightforward. While the  $(m, n)$ -channel model with its two types of subprocesses, i.e., signal and background level measurements, seemed to provide an appropriate framework for this auditory discrimination task, one can conceive of empirical situations calling for a generalization to models with more than two subsets of processes, where each subset consists of parallel processes and overall RT depends on some statistic of the subset processes.

On the practical side, the sharpness of the distribution inequality bounds should be explored for various classes of multivariate distribution functions, both with independence and with different degrees of positive or negative

dependence. Finally, in the absence of a statistical hypothesis test for some of the inequalities, numerical studies including bootstrap methods could be performed to get a better idea about the sampling variability of the bounds.

## ACKNOWLEDGEMENTS

The first author was supported by the German–American Collaborative Research Award sponsored by the American Council of Learned Societies (ACLS) and Deutscher Akademischer Austauschdienst (DAAD). The second author was supported by Postdoc Grant EL 129/1-1 from Deutsche Forschungsgemeinschaft (DFG). Thanks are due to the reviewers, Greg Francis and Duncan Luce for suggestions concerning the presentation of this work.

## REFERENCES

- Arnold, B. C., Balakrishnan, N., & Nagaraja, H. N. (1992). *A first course in order statistics*. New York: Wiley.
- Ashby, F. G., Tein, J. Y., & Balakrishnan, J. D. (1993). Response time distributions in memory scanning. *Journal of Mathematical Psychology*, *37*, 526–555.
- Bernstein, L. R., & Green, D. M. (1988). Detection of changes in spectral shape: Uniform vs. non-uniform background spectra. *Hearing Research*, *32*, 157–166.
- Bonferroni, C. E. (1937). Teoria statistica delle classi e calcolo delle probabilita. *Volume in onore di Riccardo Dalla Volta*, pp. 1–62. Florence, Italy: Universita di Firenze.
- Burbeck, S. L., & Luce, R. D. (1982). Evidence from auditory simple reaction times for both change and level detectors. *Perception & Psychophysics*, *32*, 117–133.
- Colonius, H., & Vorberg, D. (1994). Distribution inequalities for parallel models with unlimited capacity. *Journal of Mathematical Psychology*, *38*, 35–58.
- Dai, H., & Green, D. M. (1993). Discrimination of spectral shape as a function of stimulus duration. *Journal of the Acoustical Society of America*, *93*, 957–964.
- Ellermeier, W. (1996). Detectability of increments and decrements in spectral profiles. *Journal of the Acoustical Society of America*, *99*, 3119–3125.
- Ellermeier, W., & Colonius, H. (1995). Reaction times in the auditory discrimination of spectral shape. In C.-A. Possamai (Ed.), *Fechner Day 95*, pp. 185–190. Cassis, France: Intern. Soc. for Psychophysics.
- Ellermeier, W., & Colonius, H. (in preparation). Reaction times in auditory profile discrimination. *Unpublished manuscript*.
- Galambos, J. (1978). *The asymptotic theory of extreme order statistics*. New York: Wiley.
- Glaz, J. (1993). Extreme order statistics for a sequence of dependent random variables. In M. Shaked & Y. L. Tong (Eds.), *Stochastic inequalities*, IMS Lecture Notes Hayward, CA: Inst. Math. Statist. Monographs Series, Vol. 22, pp. 100–115.
- Green, D. M. (1988). *Profile analysis. Auditory intensity discrimination*. New York: Oxford University Press.
- Green, D. M. (1992). On the number of components in profile-analysis tasks. *Journal of the Acoustical Society of America*, *91*, 1616–1623.
- Kohfeld, D. L., Santee, J. L. & Wallace, N. D. (1981). Loudness and reaction time. II. Identification of detection components at different intensities and frequencies. *Perception & Psychophysics*, *29*, 550–562.
- Kounias, E. G. (1968). Bounds for the probability of a union, with applications. *Annals of Mathematical Statistics*, *39*, 2154–2158.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

- Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, **25**, 168–199.
- Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 133–168). Hillsdale, NJ: Erlbaum.
- Townsend, J. T. (1990). Serial vs parallel processing: Sometimes they look like tweedledum and tweedledee but they can (and should) be distinguished. *Psychological Science*, **1**, 46–54.
- Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.
- Townsend, J. T., & Colonius, H. (in press). Parallel processing response times and experimental determination of the stopping rule. In C. Dowling, F. Roberts & P. Theuns (Eds.), *Progress in Mathematical Psychology*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Vorberg, D. (1977). On the equivalence of parallel and serial models of information processing. Paper presented at the Tenth Annual Mathematical Psychology Meeting, Los Angeles, CA.
- Vorberg, D. (1981). Reaction time distributions predicted by serial self-terminating models of memory search. In S. Grossberg (Ed.), *Mathematical Psychology and Psychophysiology Proceedings of the Symposium in Applied Mathematics of the American Mathematical Society and the Society for Industrial and Applied Mathematics*, pp. 301–318. New York: American Mathematical Society.
- Vorberg, D., Colonius, H., & Schmidt, R. (1989). *Distribution inequalities for parallel models with unlimited capacity*. Purdue Mathematical Psychology Program Technical Report No. 89–5.
- Vorberg, D., & Ulrich, R. (1987). Random search with unequal search rates: Serial and parallel generalizations of McGill's model. *Journal of Mathematical Psychology*, **31**, 1–23.
- Worsley, K. J. (1982). An improved Bonferroni inequality and applications. *Biometrika*, **69**, 297–302.

Received: November 8, 1996