The Annals of Probability
2002, Vol. 30, No. 4, 1933–1958

DOOB, IGNATOV AND OPTIONAL SKIPPING

BY GORDON SIMONS, YI-CHING YAO AND LIJIAN YANG

University of North Carolina, Academia Sinica and Michigan State University

A general set of distribution-free conditions is described under which an i.i.d. sequence of random variables is preserved under optional skipping. This work is motivated by theorems of J. L. Doob [*Ann. of Math.* **37** (1936) 363–367] and Z. Ignatov [*Annuaire Univ. Sofia Fac. Math. Méch.* **71** (1977) 79–94], unifying and extending aspects of both.

1. Introduction and motivation. This paper discusses a general set of conditions under which an i.i.d. sequence of random variables $\xi_1, \xi_2, \dots$, taking values in a measurable space $(\mathfrak{X}, \mathfrak{B})$, with common distribution F, is preserved under "optional skipping," that is, a general set of conditions on a sequence of stopping times $\tau_1, \tau_2, \dots$, not depending on F, which guarantee that $\xi_{\tau_1}, \xi_{\tau_2}, \dots$ are i.i.d. with common distribution F. This task is motivated by fairly well-known theorems of Doob (1936) and Ignatov (1977):

I. Doob's condition simply requires the stopping times to be (a) "predictable" and (b) strictly increasing. (A stopping time τ is said to be *predictable* if the event $[\tau \le n]$ is $\sigma(\xi_1, \dots, \xi_{n-1})$-measurable for $n = 1, 2, \dots$, so that the issue of stopping at time n is independent of the value of ξ_n.)

II. Ignatov's theorem concerns real random variables and leads to the conclusion that $\xi_{\tau_1}, \xi_{\tau_2}, \dots$ are i.i.d. with common distribution F when τ_k is the occurrence time of the *first* k-record for $k = 1, 2, \dots$. [The observation ξ_τ is a *k-record* if it has rank k among the random variables $\xi_1, \dots, \xi_\tau$, that is, if the indicators $\mathbb{1}_{\{\xi_i \ge \xi_\tau\}}$, $1 \le i \le \tau$, sum to k. Ignatov's remarkable theorem asserts that the random vectors in the set $\{(\xi_{t_{k,1}}, \xi_{t_{k,2}}, \dots),\ k \ge 1\}$ are i.i.d., where $t_{k,i}$ denotes the occurrence time of the ith k-record. Various proofs and extensions of Ignatov's theorem can be found in Deheuvels (1983), Goldie and Rogers (1984), Stam (1985), Resnick (1987), Engelen, Tommassen and Vervaat (1988), Rogers (1989), Samuels (1992) and Yao (1997). Here, we are concerned with just the first components of the random vectors, that is, with the stopping times $\tau_k = t_{k,1}$ for $k = 1, 2, \dots$, and we need only observe that the first stopping time $\tau_1 = t_{1,1}$ is identically equal to 1,

Received August 2001; revised January 2002.
AMS 2000 subject classifications. 60G40, 28D05.
Key words and phrases. Ignatov's theorem, indexical stopping times, disentangled stopping times, records, k-records, optional skipping.

so that $\xi_{\tau_1}, \xi_{\tau_2}, \dots$ are i.i.d. and necessarily distributed as ξ_1, which has distribution F.]

Thus, I and II provide two rich classes of examples under which $\xi_{\tau_1}, \xi_{\tau_2}, \dots$ are i.i.d. with common distribution F. Apart from τ_1, none of the stopping times τ_k, $k \geq 2$, is predictable for case II (except in the uninteresting case of degenerate random variables $\xi_1, \xi_2, \dots$). So cases I and II provide distinct examples.

The stopping times for case II are, for the most part, unordered by index, but are "distinctly valued" (never assume a common value). In this regard, we should point out that Doob's condition (b) can be relaxed (without affecting his conclusion) to the requirement that the stopping times $\tau_1, \tau_2, \dots$ are distinctly valued. This follows from Theorem 2 in Section 3, and the discussion in I$''$ below.

III. Here is a third type of example. Let $\{B_1, B_2, \dots\} \in \mathfrak{B}$ be a countable partition of $\mathfrak{X}$, and let

$$\tau_k = \inf\{n > k : \xi_n \text{ is a member of the same partition set } B_m \text{ as } \xi_k\}, \qquad k \geq 1. \tag{1}$$

It is easily seen that the τ_k's are distinctly valued stopping times and that the stopped random variables $\xi_{\tau_1}, \xi_{\tau_2}, \dots$ are i.i.d. and distributed as $\xi_1, \xi_2, \dots$. Moreover, it is easily verified that examples of this third type are distinct from those of types I and II.

So what do these three types of examples have in common? It would seem very little, but it turns out that all of the stopping times, for all three examples, possess a common structure which guarantees that each ξ_{τ_i} $(i = 1, 2, \dots)$ is distributed as ξ_1. We call stopping times that possess this structure indexical: a stopping time τ is said to be *indexical* if, together with $A_1 = \mathfrak{X}$, there exist, for $n = 2, 3, \dots$, properly measurable nonempty random subsets $A_n = A_n(\xi_1, \dots, \xi_{n-1})$ of $\mathfrak{X}$, defined on $[\tau \geq n]$ and depending on the observations $\xi_1, \dots, \xi_{n-1}$, and countable partitions of these sets,

$$A_n = A_{n,1} + A_{n,2} + \cdots \qquad (n \geq 1;\ \text{here, } + \text{ denotes disjoint union}), \tag{2}$$

consisting of properly measurable random subsets $A_{n,k} = A_{n,k}(\xi_1, \dots, \xi_{n-1})$ $(k \geq 1)$, depending on the observations $\xi_1, \dots, \xi_{n-1}$, which satisfy the following axioms:

(3a) if $\tau \geq n$ and $\xi_n \in A_{n,1}$, then $\tau = n$;
(3b) if $\tau \geq n$ and $\xi_n \notin A_n$, then $\tau > n$ and $A_{n+1} = A_n$;
(3c) if $\tau \geq n$ and $\xi_n \in A_{n,k}$ for some $k \geq 2$, then $\tau > n$ and $A_{n+1} = A_{n,k}$.

Throughout the paper, we adopt the convention that the random variables $\xi_1, \xi_2, \dots$ are defined on the product measurable space $(\Omega, \mathfrak{F}) = (\mathfrak{X}^\infty, \mathfrak{B}^\infty)$, so that $\xi_n = \xi_n(\omega)$ is just the nth component of ω. Then P, depending on F, becomes the usual product probability measure, denoted F^∞. However, we will continue to use the standard notation P. Below, the sigma field generated by $\xi_1, \dots, \xi_n$ will be denoted by $\mathfrak{F}_n$ ($\mathfrak{F}_0 = \{\Omega, \varnothing\}$). The "proper measurability," required of A_n and $A_{n,k}$ above, takes the following form: the sets $\{(\omega, x) : \tau(\omega) \geq n,\ x \in A_n\}$ and $\{(\omega, x) : \tau(\omega) \geq n,\ x \in A_{n,k}\}$ are $\mathfrak{F}_{n-1} \times \mathfrak{B}$-measurable sets.

Observe that

$$A_{n+1} \subset A_n \qquad \text{on } [\tau > n]\ (n \geq 1), \tag{4}$$

$$\tau = \inf\{n \geq 1 : \xi_n \in A_{n,1}\} \tag{5}$$

and

$$A_{n,0} + A_{n,1} + A_{n,2} + \cdots = A_{n,0} + A_n = \mathfrak{X}, \tag{6}$$

where $A_{n,0}$ denotes the set complement of A_n.

We refer to the sets A_n, $A_{n,0}$, $A_{n,1}$ and the difference $A_n - A_{n,1}$, respectively, as the *possible set*, the *impossible set*, the *stopping set* and the *continuation set* at time n. The continuation set $A_n - A_{n,1}$ is partitioned into $A_{n,2}, A_{n,3}, \dots,$ which will be referred to as the *continuation subsets* at time n. Thus, the union of these continuation subsets is the continuation set. Notice, under axiom (3c), that the points of the continuation subset $A_{n,j}$ $(j \geq 2)$ become impossible points at time $n+1$ if $\xi_n \in A_{n,k}$ for some $k \geq 2$, $k \neq j$.

It is shown in Section 2 (Theorem 1), when τ is indexical, that the conditional distribution of ξ_τ given $\mathfrak{F}_{n-1}$ assumes the simple form

$$P(\xi_\tau \in B | \xi_1, \dots, \xi_{n-1}) = F(B|A_n) := \frac{F(B \cap A_n)}{F(A_n)} \qquad \text{on the set } [\tau \geq n]. \tag{7}$$

This holds for $n = 1, 2, \dots,$ whenever the distribution F is such that $P(\tau < \infty) = 1$. Since $A_1 = \mathfrak{X}$, it follows trivially from (7), when $n = 1$, that $P(\xi_\tau \in B) = F(B)$, $B \in \mathfrak{B}$. So, as required, ξ_τ has distribution F (that of ξ_1) whenever F and indexical τ are such that $P(\tau < \infty) = 1$.

Remarkably, under considerable generality, in order for condition (7) to hold for $n \geq 1$, and for all distributions F, the stopping time τ must be indexical. That is, from (7), one can uncover the partitions of the sets A_n, $n \geq 1$, appearing in (2), and to show that the indexicality axioms (3a)–(3c) must hold. The precise statement of this converse to Theorem 1 (Proposition 1) is stated and proven in Section 4.

We next describe the stopping times appearing in examples I, II and III as indexical stopping times:

I′. Each predictable stopping time τ of type I has the possible set $A_n = \mathfrak{X}$ for every n, and stopping set $A_{n,1} = \mathfrak{X}$, or $= \varnothing$ (the empty set), depending on whether or not $\tau = n$ (this depending on just the random variables $\xi_1, \dots, \xi_{n-1}$). The partition of A_n consists of the stopping set $A_{n,1}$ and the single continuation subset $A_{n,2} = \mathfrak{X} - A_{n,1}$, either $\varnothing$ or $\mathfrak{X}$, again depending on whether or not $\tau = n$. This description meets the requirements for indexicality. So predictable stopping times are indexical.

II′. The "Ignatov stopping times" described in II are only slightly more complicated. τ_1 is identically equal to 1 and qualifies as a predictable stopping time; nothing more needs to be said about it. For τ_k with $k \geq 2$, the possible set A_n consists of those points x in $\mathfrak{X}$ that have rank less than or equal to k among $\{\xi_1, \ldots, \xi_{n-1}, x\}$. Those points $x \in A_n$ for which the rank is k go into the stopping set $A_{n,1}$, and the remainder go into the single continuation subset $A_{n,2}$. This description meets the requirements for indexicality. So these Ignatov stopping times are indexical. The "nesting" of the possible sets A_n [see (4) above] is apparent. Since there must be a finite first k-record, a.s., for every k, independent of the distribution F, all of the stopping times for this example are a.s. finite. Note that if no x satisfies $F(x-) < F(x) = 1$, then all ith k-records exist with probability 1, in which case only the occurrence time of the first k-record is indexical; the remaining k-records are stochastically larger than ξ_1. If $F(x-) < F(x) = 1$ occurs for some x, then no ith k-record ($i \geq 2$) exists with a positive probability.

III′. Since $\tau_k > k$ in example III, one needs a single continuation set $A_{n,2} = \mathfrak{X}$ and an empty stopping set $A_{n,1}$ when $n < k$, so that $A_n = \mathfrak{X}$. This keeps sampling going through the kth observation. For $n = k$, one must use the sets $A_{k,i}$, $i = 2, 3, \ldots$, to store information about the partition elements $\{B_m : m = 1, 2, \ldots\}$. It is convenient to set $A_{k,i} = B_{i-1}$ for $i = 2, 3, \ldots$ ($A_k = \sum_{m \geq 1} B_m = \mathfrak{X}$ and $A_{k,1} = \varnothing$). Then, for $n > k$, if ξ_k assumes a value within the partition element B_m, one sets $A_n = A_{n,1} = B_m$ and $A_{n,0} = B_m^c$, and waits for the next observation to occur in B_m. The resulting stopping time τ_k is indexical. Clearly, all of the τ_k's are finite a.s., independent of the choice of distribution F.

So indexical stopping times preserve the common distribution of the i.i.d. sequence $\xi_1, \xi_2, \ldots$. What preserves the independence? We turn to this issue next. A pair of indexical stopping times τ_1 and τ_2, arising from the indexical sets $A_n^{(1)}$, $A_{n,k}^{(1)}$ and $A_n^{(2)}$, $A_{n,k}^{(2)}$, respectively, will be called *disentangled* if for each $n = 1, 2, \ldots$, and each ω-point in $[\tau_1 \geq n,\ \tau_2 \geq n]$, there exists an index $s = s(\omega) \in \{1, 2\}$ such that one of the following statements holds:

(8a) $A_n^{(s)} \subset A_{n,k}^{(3-s)}$ for some $k \neq 1$ (impossible set or single continuation subset on the right);

(8b) $A_n^{(s)} = A_{n,2}^{(s)}$ (possible set for index s consists of exactly one continuation subset);

(8c) $\operatorname{card}(A_n^{(s)}) = 1$ (possible set for index s consists of a single point).

While (8b) may be replaced by $A_n^{(s)} = A_{n,k}^{(s)}$ for some $k \geq 2$, we shall adopt the convention, when there is only one continuation subset, of denoting it by $A_{n,2}$ (superscripted as needed).

It turns out, whenever the indexical stopping times τ_1 and τ_2 are disentangled, and F is such that these stopping times are almost surely finite, that the stopped random variables ξ_{τ_1} and ξ_{τ_2} are independent. Surprisingly, to extend this independence to the entire sequence $\xi_{\tau_1}, \xi_{\tau_2}, \dots$, it is only necessary to assume that the indexical stopping times $\tau_1, \tau_2, \dots$ are *pairwise* disentangled. That is, it is only necessary to verify disentanglement for each distinct pair of (indexical) stopping times. A precise statement to this effect, with proof, is given in Section 3 (Theorem 2). A converse to Theorem 2 appears in Section 4 (Proposition 2), in which the naturalness of conditions (8a)–(8c) is revealed.

We next describe the stopping times appearing in examples I, II and III as pairwise disentangled stopping times:

I$''$. For distinct-valued predictable (indexical) stopping times $\tau^{(1)}$ and $\tau^{(2)}$, we have $A_n^{(1)} = A_n^{(2)} = \mathfrak{X}$ for every n (see I$'$). In examining what happens for $[\tau^{(1)} \geq n,\ \tau^{(2)} \geq n]$, we have three cases to consider: (i) if $[\tau^{(1)} = n,\ \tau^{(2)} > n]$, then condition (8b) obtains with $s = 2$: $A_n^{(2)} = A_{n,2}^{(2)} = \mathfrak{X}$ (see I$'$ above). Moreover, condition (8a) obtains with $s = 1$: $\mathfrak{X} = A_n^{(1)} \subset A_{n,2}^{(2)} = \mathfrak{X}$; (ii) if $[\tau^{(1)} > n,\ \tau^{(2)} = n]$, the situation is like that of case (i) with the indices reversed; (iii) if $[\tau^{(1)} > n,\ \tau^{(2)} > n]$, then $A_n^{(1)} = A_n^{(2)} = A_{n,2}^{(1)} = A_{n,2}^{(2)} = \mathfrak{X}$, so that conditions (8a) and (8b) obtain with $s = 1$ and 2. So distinctly valued predictable stopping times are pairwise disentangled.

II$''$. Let $\tau^{(1)} = t_{k,1}$ and $\tau^{(2)} = t_{l,1}$ denote the occurrence times of the first k-record and first l-record, respectively, and suppose, for definiteness, that $k < l$. For fixed n and $x \in \mathfrak{X}$, let $r(x) := 1 + \sum_{i=1}^{n-1} \mathbb{1}_{\{\xi_i \geq x\}}$ denote the rank of x among $\{\xi_1, \dots, \xi_{n-1}, x\}$. Then, on $[\tau^{(1)} \geq n,\ \tau^{(2)} \geq n] = [t_{k,1} \geq n, t_{l,1} \geq n]$, we have $A_n^{(1)} = \{x : r(x) \leq k\} \subset \{x : r(x) < l\} = A_{n,2}^{(2)}$ (see II$'$). Consequently, condition (8a) always obtains with $s = 1$. So the occurrence times of first k-records, $k = 1, 2, \dots$, are pairwise disentangled. It should be remarked that, while Ignatov's theorem implies the independence of the corresponding stopped random variables, Theorem 2 below says that these stopped random variables are indeed (conditionally) independent given $\xi_1, \dots, \xi_{n-1}$ for each $n = 1, 2, \dots$ ($n = 1$ corresponding to Ignatov's independence result).

III$''$. Let $\tau^{(1)} = \tau_k$ and $\tau^{(2)} = \tau_l$, as defined in (1), with $k < l$, and suppose for some fixed n and ω that $[\tau^{(1)} \geq n,\ \tau^{(2)} \geq n]$. If $n < l$, then condition (8b) holds with $s = 2$: $A_n^{(2)} = A_{n,2}^{(2)} = \mathfrak{X}$ (see III$'$). When $n = l$ ($> k$), condition (8a) holds with $s = 1$ and $k = m^{(1)} + 1$: $A_n^{(1)} = A_{n,1}^{(1)} = B_{m^{(1)}} = A_{n,(m^{(1)}+1)}^{(2)}$ (see III$'$), where $B_{m^{(1)}}$ is the partition element B_m that contains ξ_k. Finally, when $n > l$, we note that the partition element $B_{m^{(2)}}$ containing ξ_l must be disjoint from $B_{m^{(1)}}$, so that condition (8a) holds for $s = 1$ and 2: $A_n^{(s)} = B_{m^{(s)}} \subset B_{m^{(3-s)}}^c = A_{n,0}^{(3-s)}$ for $s = 1$ and 2 (see III$'$). Consequently, the collection of stopping times described in III are pairwise disentangled.

Despite the strong suggestion, provided by the examples of types I, II and III, that indexical stopping times must be distinctly valued to obtain the independence of the corresponding stopped random variables, such a requirement is neither sufficient nor necessary. We can illustrate this point for $\mathfrak{X} = \{0, 1\}$ with pairs of indexical stopping times. Let $\tau \equiv 1$, $\sigma = \inf\{n \geq 2 : \xi_n = \xi_1\}$, and $\gamma \equiv 2$, all indexical. The first two are distinctly valued, but, obviously, $P(\xi_\sigma = \xi_\tau) = 1$: no independence. The second and third stopping times can assume the common value 2, and, yet, ξ_σ and ξ_γ are independent. The two indexical stopping times σ and γ are disentangled; they illustrate the need for condition (8c).

We close this section with an example of a nonindexical stopping time τ for which $\xi_\tau \overset{\mathcal{D}}{=} \xi_1$ (equality in distribution) for *every* distribution function F. Let $\xi_1, \xi_2, \dots$ be i.i.d. Bernoulli random variables taking values in $\mathfrak{X} = \{0, 1\}$, with $\mathfrak{B}$ its power set. Here, the entire class of possible distributions F is parameterized by $p = P(\xi_1 = 1) \in [0, 1]$. Consider the stopping time τ which stops at the right endpoint of the following complete set of nine sample paths

$$(\mathbf{0}, \mathbf{1}),\ (0, 0, 0),\ (1, 0, 0),\ (1, 1, 0),\ (\mathbf{1}, \mathbf{1}, \mathbf{1}),$$
$$(0, 0, 1, 0),\ (\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{1}),\ (1, 0, 1, 0),\ (\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{1}),$$

occurring, respectively, with probabilities

$$\mathbf{p}(\mathbf{1} - \mathbf{p}),\ (1 - p)^3,\ p(1 - p)^2,\ p^2(1 - p),\ \mathbf{p}^{\mathbf{3}},$$
$$p(1 - p)^3,\ \mathbf{p}^{\mathbf{2}}(\mathbf{1} - \mathbf{p})^{\mathbf{2}},\ p^2(1 - p)^2,\ \mathbf{p}^{\mathbf{3}}(\mathbf{1} - \mathbf{p}),$$

adding to 1. (Every path is attainable, and the set is exhaustive.) Thus, the range of τ is $\{2, 3, 4\}$. The first, fifth, seventh and ninth paths, appearing in bold type (as do the corresponding probabilities), result in $\xi_\tau = 1$, and this occurs with probability $p(1 - p) + p^3 + p^2(1 - p)^2 + p^3(1 - p) = p$. If τ were indexical, then, clearly, the possible set A_2 when $\xi_1 = 0$ would have to be all of $\mathfrak{X}$, so that, according to equation (7) and Theorem 1, we would have $P(\xi_\tau = 1 | \xi_1 = 0) = F(\{1\})/F(\mathfrak{X}) = p/1 = p$. However, by direct calculation,

$$\begin{aligned} P(\xi_\tau = 1 | \xi_1 = 0) &= \frac{P(\xi_\tau = 1,\ \xi_1 = 0)}{P(\xi_1 = 0)} \\ &= \frac{p(1 - p) + p^2(1 - p)^2}{1 - p} = p + p^2(1 - p) > p \end{aligned}$$

for $0 < p < 1$. So τ is not indexical. This can also be shown, more directly, from the definition.

2. Establishing the right distribution for the stopped r.v.

The following theorem validates the use of (7) as described in the Introduction.

THEOREM 1. *Let $\xi_1, \xi_2, \dots$ be i.i.d. random variables with common distribution F on $(\mathfrak{X}, \mathfrak{B})$. Further, let τ be an indexical stopping time, with associated indexical sets A_n, $A_{n,i}$, for which $P(\tau < \infty) = 1$. Then, for each $n = 1, 2, \dots$,*

$$Q_n(B, \omega) := \frac{F(B \cap A_n)}{F(A_n)} = \frac{\int_B \mathbb{1}_{A_n}(x) F(dx)}{\int_{\mathfrak{X}} \mathbb{1}_{A_n}(x) F(dx)} \qquad (\textit{here, } 0/0 \textit{ is understood as } 0) \tag{9}$$

defines a regular version of the conditional probability $P(\xi_\tau \in B | \mathfrak{F}_{n-1})$ on $\mathfrak{B} \times [\omega : \tau \geq n]$.

PROOF. It must be shown that (i) $Q_n(\cdot, \omega) = \frac{F(\cdot \cap A_n)}{F(A_n)}$ is a probability measure on $(\mathfrak{X}, \mathfrak{B})$, for almost all ω for which $\tau = \tau(\omega) \geq n$, and (ii) $Q_n(B, \cdot) = \frac{F(B \cap A_n(\cdot))}{F(A_n(\cdot))}$ is a version of the conditional probability $P(\xi_\tau \in B | \xi_1, \dots, \xi_{n-1})$ on $[\tau \geq n]$ for each fixed $B \in \mathfrak{B}$. Clearly, (i) is obvious providing $F(A_n) > 0$ a.s. The proof of (ii) requires truncation and induction. The fact that $F(A_n) > 0$ a.s. on $[\tau \geq n]$ is indeed a consequence of (ii) since $P(\xi_\tau \in \mathfrak{X} | \mathfrak{F}_{n-1}) = F(\mathfrak{F} \cap A_n)/F(A_n) = 0/0 = 0$ on $\Omega' = \{\omega : \tau \geq n,\ F(A_n) = 0\} \in \mathfrak{F}_{n-1}$, from which it follows that $P(\Omega') = 0$. So we need to establish (ii).

Informally stated, the stopping time τ is truncated at time r by enforcing stopping at the first time $n \geq r$ for which ξ_n is *contained in a possible set*. Stopping any sooner would destroy the indexicality possessed by τ, which must be preserved. Formally, the stopping time $\tau^{(r)}$ is the *indexical truncation of τ at time r* if it stops according to τ before time r and stops thereafter at the first time n for which $\xi_n \in A_r$. According to this prescription, the analogues of A_n and $A_{n,i}$, appropriate to $\tau^{(r)}$, remain unchanged before time r. For time $n \geq r$, the possible set assumes the form $A_n^{(r)} = A_r$, which, in turn is "partitioned" into a stopping set and no continuation subsets: $A_{n,1}^{(r)} = A_n^{(r)} = A_r$. It is apparent from this description that $\tau^{(r)}$ is conditionally distributed as a geometric random variable when $n \geq r$. The analogue of $Q_n(B, \omega)$, appropriate to $\tau^{(r)}$, assumes the form

$$Q_n^{(r)}(B, \omega) = \frac{F(B \cap A_{n \wedge r})}{F(A_{n \wedge r})} \qquad \text{on } [\tau^{(r)} \geq n], \tag{10}$$

where $n \wedge r = \min(n, r)$.

We claim that $P(\xi_{\tau^{(r)}} \in B | \mathfrak{F}_{n-1})$, on $[\tau^{(r)} \geq n]$, is a.s. as described in (10). Assume, for the moment, that the claim is true, and observe that

$$\begin{aligned} |P(\xi_{\tau^{(r)}} \in B | \mathfrak{F}_{n-1}) - P(\xi_\tau \in B | \mathfrak{F}_{n-1})| &= \big|E\big[\big(\mathbb{1}_{\{\xi_{\tau^{(r)}} \in B\}} - \mathbb{1}_{\{\xi_\tau \in B\}}\big) \big| \mathfrak{F}_{n-1}\big]\big| \\ &\leq E\big[\big|\mathbb{1}_{\{\xi_{\tau^{(r)}} \in B\}} - \mathbb{1}_{\{\xi_\tau \in B\}}\big| \big| \mathfrak{F}_{n-1}\big] \\ &\leq E\big[\mathbb{1}_{\{\tau \geq r\}} \big| \mathfrak{F}_{n-1}\big] \\ &\to 0 \qquad \text{a.s. as } r \to \infty. \end{aligned}$$

So $P(\xi_{\tau^{(r)}} \in B|\mathfrak{F}_{n-1}) \to P(\xi_\tau \in B|\mathfrak{F}_{n-1})$ as $r \to \infty$. Hence, by the claim and noting that $[\tau^{(r)} \geq n] = [\tau \geq n]$ for $r \geq n$, we have

$$\begin{aligned} P(\xi_\tau \in B|\mathfrak{F}_{n-1}) &= \lim_{r\to\infty} \frac{F(B \cap A_{n\wedge r})}{F(A_{n\wedge r})} \\ &= \frac{F(B \cap A_n)}{F(A_n)} = Q_n(B, \omega) \qquad \text{a.s. on } [\tau \geq n]. \end{aligned}$$

So it remains to establish the claim. This will require backward induction.

For $n \geq r$, on $[\tau^{(r)} \geq n]$, $\tau^{(r)}$ stops at the first time $k \geq n$ for which $\xi_k \in A_r$. Thus, on $[\tau^{(r)} \geq n]$,

$$P(\xi_{\tau^{(r)}} \in B|\mathfrak{F}_{n-1}) = \frac{F(B \cap A_r)}{F(A_r)} = \frac{F(B \cap A_{n\wedge r})}{F(A_{n\wedge r})} = Q_n^{(r)}(B, \omega) \qquad \text{a.s.,}$$

establishing the claim for $n \geq r$.

We proceed by backward induction on $n = r - 1, r - 2, \ldots, 1$. Suppose (10) holds for the stopping time $\tau^{(r)}$ when n is replaced by $n + 1 \leq r$, that is, that on $[\tau^{(r)} \geq n + 1]$,

$$P(\xi_{\tau^{(r)}} \in B|\mathfrak{F}_n) = Q_{n+1}^{(r)}(B, \omega) = \frac{F(B \cap A_{n+1})}{F(A_{n+1})} \qquad \text{a.s.}$$

Note that $A_r^{(r)} = A_r$ and $A_{n,i}^{(r)} = A_{n,i}$ for $n < r$, so that the superscript "(r)" in $A_{n,i}^{(r)}$ will be suppressed below.

On $[\tau^{(r)} \geq n]$,

$$\begin{aligned} P(\xi_{\tau^{(r)}} \in B|\mathfrak{F}_{n-1}) &= E\big(\mathbb{1}_{\{\xi_{\tau^{(r)}} \in B\}}|\mathfrak{F}_{n-1}\big) \\ &= E\left\{E\left[\sum_{i=0}^{\infty} \mathbb{1}_{\{\xi_{\tau^{(r)}} \in B,\ \xi_n \in A_{n,i}\}}\Big|\mathfrak{F}_n\right]\Big|\mathfrak{F}_{n-1}\right\} \\ &= \sum_{i=0}^{\infty} E\Big\{E\big[\mathbb{1}_{\{\xi_{\tau^{(r)}} \in B,\ \xi_n \in A_{n,i}\}}\big|\mathfrak{F}_n\big]\big|\mathfrak{F}_{n-1}\Big\} \\ &= \sum_{i=0}^{\infty} E\{T_i|\mathfrak{F}_{n-1}\}, \end{aligned} \tag{11}$$

where $T_i = E(\mathbb{1}_{\{\xi_{\tau^{(r)}} \in B,\ \xi_n \in A_{n,i}\}}|\mathfrak{F}_n)$ for $i = 0, 1, \ldots$. When, $\xi_n \in A_{n,0}$ (the complement of A_n), axiom (3b) tells us that $[\tau^{(r)} \geq n + 1]$ and $A_{n+1} = A_n$, so that

$$\begin{aligned} T_0 &= E\big(\mathbb{1}_{\{\xi_{\tau^{(r)}} \in B,\ \xi_n \in A_{n,0}\}}\big|\mathfrak{F}_n\big) \\ &= \mathbb{1}_{\{\xi_n \in A_{n,0}\}} P(\xi_{\tau^{(r)}} \in B|\mathfrak{F}_n) \\ &= \mathbb{1}_{\{\xi_n \in A_{n,0}\}} \frac{F(B \cap A_{n+1})}{F(A_{n+1})} \qquad \text{(by the induction hypothesis)} \\ &= \mathbb{1}_{\{\xi_n \in A_{n,0}\}} \frac{F(B \cap A_n)}{F(A_n)}. \end{aligned} \tag{12}$$

When $\xi_n \in A_{n,1}$, axiom (3a) tells us that $[\tau^{(r)} = n]$, so that

$$T_1 = E\big(\mathbb{1}_{\{\xi_n \in B \cap A_{n,1}\}} \big| \mathfrak{F}_n\big) = \mathbb{1}_{\{\xi_n \in B \cap A_{n,1}\}}. \tag{13}$$

When $\xi_n \in A_{n,i}$ with $i \geq 2$, axiom (3c) tells us that $[\tau^{(r)} \geq n+1]$ and $A_{n+1} = A_{n,i}$, so that

$$\begin{aligned} T_i &= E\big(\mathbb{1}_{\{\xi_{\tau^{(r)}} \in B,\ \xi_n \in A_{n,i}\}} \big| \mathfrak{F}_n\big) \\ &= \mathbb{1}_{\{\xi_n \in A_{n,i}\}} P(\xi_{\tau^{(r)}} \in B | \mathfrak{F}_n) \\ &= \mathbb{1}_{\{\xi_n \in A_{n,i}\}} \frac{F(B \cap A_{n+1})}{F(A_{n+1})} \qquad \text{(by the induction hypothesis)} \\ &= \mathbb{1}_{\{\xi_n \in A_{n,i}\}} \frac{F(B \cap A_{n,i})}{F(A_{n,i})}. \end{aligned} \tag{14}$$

Combining (11)–(14), and noting that $A_n = \sum_{i=1}^{\infty} A_{n,i}$, we obtain

$$\begin{aligned} P\big(\xi_{\tau^{(r)}} \in B \big| \mathfrak{F}_{n-1}\big) &= \sum_{i=0}^{\infty} E\{T_i | \mathfrak{F}_{n-1}\} \\ &= \frac{F(B \cap A_n)}{F(A_n)} F(A_{n,0}) + F(B \cap A_{n,1}) \\ &\quad + \sum_{i=2}^{\infty} \frac{F(B \cap A_{n,i})}{F(A_{n,i})} F(A_{n,i}) \\ &= \frac{F(B \cap A_n)}{F(A_n)} \big(1 - F(A_n)\big) + F(B \cap A_n) \\ &= \frac{F(B \cap A_n)}{F(A_n)} = Q_n^{(r)}(B, \omega). \end{aligned}$$

This completes the induction step, and the proof. □

3. Independence of the stopped sequence. For this section, we assume that $\mathfrak{X}$ is a complete separable metric space (a Polish space) with $\mathfrak{B}$ the corresponding σ-field of Borel sets. We are uncertain whether this restriction is genuinely needed at this point, but we are unable to carry out an induction step in the proof of our next theorem without it. In any event, this represents a modest restriction on our sequence of random variables $\xi_1, \xi_2, \ldots$.

THEOREM 2. *For every distribution F on $(\mathfrak{X}, \mathfrak{B})$ and any collection $\mathfrak{T}$ of pairwise disentangled indexical stopping times τ with $\tau < \infty$ a.s., the corresponding stopped random variables ξ_τ, $\tau \in \mathfrak{T}$, are conditionally independent given $\mathfrak{F}_n$ for $n = 0, 1, \ldots$.*

The reader is reminded that, by definition, the σ-field $\mathfrak{F}_0$ is degenerate. Thus this theorem states, when $n = 0$, that the stopped random variables ξ_τ, $\tau \in \mathfrak{T}$, for which $\tau < \infty$ a.s. are (unconditionally) independent. Moreover, since they are indexical, Theorem 1 asserts that these stopped random variables have common distribution F.

PROOF OF THEOREM 2. The heart of our proof is backward induction: if the stopped random variables ξ_τ with τ a.s. finite are conditionally independent given $\mathfrak{F}_n$, then, under the assumptions of indexicality and pairwise disentanglement, they are conditionally independent given $\mathfrak{F}_{n-1}$. Beyond this, an appropriate truncation argument is needed to get the backward induction started.

The induction step. Fix F. It is enough to work with a finite but variable number of a.s. finite stopping times $\tau_1, \dots, \tau_r$ in $\mathfrak{T}$. If

$$\begin{aligned} &P(\xi_{\tau_i} \in B_i, i = 1, \dots, r | \mathfrak{F}_n) \\ &\quad = \prod_{i=1}^{r} P(\xi_{\tau_i} \in B_i | \mathfrak{F}_n) \qquad \text{for all } B_1, \dots, B_r \text{ in } \mathfrak{B} \text{ and } r = 1, 2, \dots, \end{aligned} \tag{15}$$

for some $n \geq 1$, then we must show that

$$\begin{aligned} &P(\xi_{\tau_i} \in B_i, i = 1, \dots, r | \mathfrak{F}_{n-1}) \\ &\quad = \prod_{i=1}^{r} P(\xi_{\tau_i} \in B_i | \mathfrak{F}_{n-1}) \qquad \text{for all } B_1, \dots, B_r \text{ in } \mathfrak{B} \text{ and } r = 1, 2, \dots. \end{aligned} \tag{16}$$

Our proof of this will include the use of forward induction based on r. Of course, (16) holds for $r = 1$. Thus the induction hypothesis assumes the truth of (15) and the truth of the equality in (16) for up to $r - 1$ stopping times ($r \geq 2$), and the task is to show the equality in (16) for r stopping times.

To begin with, we observe, for a given r, when one of the stopping times τ_j in (16) is less than n that, given $\mathfrak{F}_{n-1}$, one knows whether or not ξ_{τ_j} is in B_j. If it is not in B_j, both sides of (16) assume the value zero, and there is nothing more to prove. If it is in B_j, then the index j can be removed from both sides of (16) and the task that remains is to establish the validity of

$$P(\xi_{\tau_i} \in B_i, i = 1, \dots, r, \ i \neq j | \mathfrak{F}_{n-1}) = \prod_{i=1,\ i \neq j}^{r} P(\xi_{\tau_i} \in B_i | \mathfrak{F}_{n-1}),$$

which involves $r - 1$ stopping times. Then the induction hypothesis for $r - 1$ establishes this. What is left to establish is the equality in (16), for the given r, on the set $[\tau_i \geq n, \text{ for } i = 1, \dots, r]$.

As a second step, we recall, from Theorem 1, that

$$P(\xi_{\tau_i} \in B_i | \mathfrak{F}_{n-1}) = \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})} \qquad \text{on } [\tau_i \geq n], \tag{17}$$

where $A_n^{(i)}$ refers to the possible set at time n for the indexical stopping time τ_i. (In addition, we shall have need to refer to the impossible set $A_{n,0}^{(i)}$, the stopping set $A_{n,1}^{(i)}$ and the continuation subsets $A_{n,k}^{(i)}$, $k \geq 2$, for τ_i at time n.) A small technical point: we established, within the body of the proof of Theorem 1, that the denominator on the right-hand side of (17) is strictly positive a.s. Consequently, what must be established for all $B_1, \dots, B_r$ in $\mathfrak{B}$ is the formula

$$\begin{aligned} &P(\xi_{\tau_i} \in B_i, i = 1, \dots, r | \mathfrak{F}_{n-1}) \\ &\quad = \prod_{i=1}^{r} \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})} \qquad \text{on } [\tau_i \geq n, \text{ for } i = 1, \dots, r]. \end{aligned} \tag{18}$$

The third step is to partition the set $T = [\tau_i \geq n, \text{ for } i = 1, \dots, r]$ into subsets based on the conditions (8a)–(8c) appearing in the definition of disentangled stopping times. It is enough to work with (8a)–(8c) for a single pair of stopping times. For definiteness, we will work with the first pair. Then the task will be to establish (18) on each of these subsets. Since the conditions (8a)–(8c) are not always mutually exclusive, the subsets we describe might actually overlap, but this will not matter; their union will cover T.

Corresponding to (8a), we must introduce the subsets

$$\Omega(s,k) := \{\omega \in T : A_n^{(s)} \subset A_{n,k}^{(3-s)}\}, \qquad s = 1, 2, \ k \neq 1. \tag{19}$$

Here, k assumes the values $0, 2, 3, \dots$. The reader is reminded that the expression to the right of "T:" does depend on ω through the values of the observations $\xi_1, \dots, \xi_{n-1}$, which are suppressed in the notation. Corresponding to (8b), we must introduce two more subsets:

$$\Omega(s)' := \{\omega \in T : A_n^{(s)} = A_{n,2}^{(s)}\}, \qquad s = 1, 2. \tag{20}$$

Corresponding to (8c) we need two additional sets:

$$\Omega(s)'' := \{\omega \in T : \operatorname{card}(A_n^{(s)}) = 1\}, \qquad s = 1, 2. \tag{21}$$

It follows from the disentanglement of τ_1 and τ_2 that the union of all of these subsets is the $\mathfrak{F}_{n-1}$-measurable subset T.

Now these subsets of T might not be $\mathfrak{F}_{n-1}$-measurable, a technical difficulty which, potentially, could derail our proof of the induction step. However, we claim that all of them are "universally measurable sets" in the sense that they belong to the completion of $\mathfrak{F}_{n-1}$, under F, for every distribution function F on $(\mathfrak{X}, \mathfrak{B})$; this is adequate for our purposes. The proof of this claim appears in the Appendix. (This is the only place that we require the topological assumption that we made at the start of this section.)

Establishing (18) *on* $\Omega(s,k)$. For definiteness, we take s to be 1. On $\Omega(1,k)$, we have $A_n^{(1)} \subset A_{n,k}^{(2)}$, where $k = 0$ or $k \geq 2$. If, in addition, $\xi_n \in A_n^{(1)}$, then it follows from (3b) and (3c) that $A_{n+1}^{(2)} = A_{n,*}^{(2)}$, where $A_{n,*}^{(2)}$ denotes $A_n^{(2)}$ or $A_{n,k}^{(2)}$ as $k = 0$ or $k \ge 2$, respectively. Consequently,

$$P(\xi_{\tau_2} \in B_2 | \mathfrak{F}_n) = \frac{F(B_2 \cap A_{n+1}^{(2)})}{F(A_{n+1}^{(2)})} = \frac{F(B_2 \cap A_{n,*}^{(2)})}{F(A_{n,*}^{(2)})}, \tag{22}$$

where the rightmost side is $\mathfrak{F}_{n-1}$-measurable. On the other hand, if $\xi_n \notin A_n^{(1)}$, then it follows from (3b) that $A_{n+1}^{(1)} = A_n^{(1)}$, so that

$$P(\xi_{\tau_1} \in B_1 | \mathfrak{F}_n) = \frac{F(B_1 \cap A_{n+1}^{(1)})}{F(A_{n+1}^{(1)})} = \frac{F(B_1 \cap A_n^{(1)})}{F(A_n^{(1)})}, \tag{23}$$

where the rightmost side is $\mathfrak{F}_{n-1}$-measurable.

Now

$$\begin{aligned} P(\xi_{\tau_i} \in B_i, i = 1, \dots, r | \mathfrak{F}_{n-1}) &= E\big(P(\xi_{\tau_i} \in B_i, i = 1, \dots, r | \mathfrak{F}_n) \big| \mathfrak{F}_{n-1}\big) \\ &= K_1 + K_2, \end{aligned}$$

where

$$K_1 = E\Big(P(\xi_{\tau_i} \in B_i, i = 1, \dots, r | \mathfrak{F}_n) \cdot \mathbb{1}_{\{\xi_n \in A_n^{(1)}\}} \big| \mathfrak{F}_{n-1}\Big)$$

and

$$K_2 = E\Big(P(\xi_{\tau_i} \in B_i, i = 1, \dots, r | \mathfrak{F}_n) \cdot \mathbb{1}_{\{\xi_n \notin A_n^{(1)}\}} \big| \mathfrak{F}_{n-1}\Big).$$

From (22), then (23), and the induction assumption, we obtain

$$\begin{aligned} K_1 &= \frac{F(B_2 \cap A_{n,*}^{(2)})}{F(A_{n,*}^{(2)})} E\Big(P(\xi_{\tau_i} \in B_i, i \ne 2 | \mathfrak{F}_n) \cdot \mathbb{1}_{\{\xi_n \in A_n^{(1)}\}} \big| \mathfrak{F}_{n-1}\Big) \\ &= \frac{F(B_2 \cap A_{n,*}^{(2)})}{F(A_{n,*}^{(2)})} E\Big(P(\xi_{\tau_i} \in B_i, i \ne 2 | \mathfrak{F}_n)\big(1 - \mathbb{1}_{\{\xi_n \notin A_n^{(1)}\}}\big) \big| \mathfrak{F}_{n-1}\Big) \\ &= \frac{F(B_2 \cap A_{n,*}^{(2)})}{F(A_{n,*}^{(2)})} \\ &\quad \times \Bigg[\prod_{\substack{i=1 \\ i \ne 2}}^{r} \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})} - \frac{F(B_1 \cap A_n^{(1)})}{F(A_n^{(1)})} \\ &\qquad\qquad \times E\Big(P(\xi_{\tau_i} \in B_i, i = 3, \dots, r | \mathfrak{F}_n) \cdot \mathbb{1}_{\{\xi_n \notin A_n^{(1)}\}} \big| \mathfrak{F}_{n-1}\Big)\Bigg], \end{aligned}$$

and from (23), then (22), and the induction assumption, we obtain

$$\begin{aligned}
K_2 &= \frac{F(B_1 \cap A_n^{(1)})}{F(A_n^{(1)})} E\Big(P(\xi_{\tau_i} \in B_i, i = 2, \ldots, r | \mathfrak{F}_n) \cdot \mathbb{1}_{\{\xi_n \notin A_n^{(1)}\}} \big| \mathfrak{F}_{n-1}\Big) \\
&= \frac{F(B_1 \cap A_n^{(1)})}{F(A_n^{(1)})} E\Big(P(\xi_{\tau_i} \in B_i, i = 2, \ldots, r | \mathfrak{F}_n)\Big(1 - \mathbb{1}_{\{\xi_n \in A_n^{(1)}\}}\Big) \big| \mathfrak{F}_{n-1}\Big) \\
&= \frac{F(B_1 \cap A_n^{(1)})}{F(A_n^{(1)})} \\
&\quad \times \Bigg[\prod_{i=2}^{r} \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})} - \frac{F(B_2 \cap A_{n,*}^{(2)})}{F(A_{n,*}^{(2)})} \\
&\qquad \times E\Big(P(\xi_{\tau_i} \in B_i, i = 3, \ldots, r | \mathfrak{F}_n) \cdot \mathbb{1}_{\{\xi_n \in A_n^{(1)}\}} \big| \mathfrak{F}_{n-1}\Big)\Bigg].
\end{aligned}$$

Adding K_1 and K_2, combining the last terms of K_1 and K_2 and then using the induction assumption once more, we obtain

$$\begin{aligned}
&\frac{F(B_2 \cap A_{n,*}^{(2)})}{F(A_{n,*}^{(2)})} \prod_{\substack{i=1 \\ i \neq 2}}^{r} \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})} \\
&\quad + \prod_{i=1}^{r} \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})} - \frac{F(B_2 \cap A_{n,*}^{(2)})}{F(A_{n,*}^{(2)})} \prod_{\substack{i=1 \\ i \neq 2}}^{r} \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})} \\
&\quad = \prod_{i=1}^{r} \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})}.
\end{aligned}$$

This completes the proof for ω-points in subsets of the form $\Omega(s, k)$.

Establishing (18) *on* $\Omega'(s)$. For definiteness, we take s to be 1. On $\Omega'(1)$, if $\xi_n \in A_n^{(1)} = A_{n,2}^{(1)}$, then it follows by (3c) that $A_{n+1}^{(1)} = A_{n,2}^{(1)}$ $(= A_n^{(1)})$. If, instead, $\xi_n \notin A_n^{(1)}$, we must have $A_{n+1}^{(1)} = A_n^{(1)}$ by (3b). So for ω in $\Omega'(1)$, we have $A_{n+1}^{(1)} = A_n^{(1)}$, so that (23) holds. Thus, by the induction assumption,

$$\begin{aligned}
&P(\xi_{\tau_i} \in B_i, i = 1, \ldots, r | \mathfrak{F}_{n-1}) \\
&\quad = E\big(P(\xi_{\tau_i} \in B_i, i = 1, \ldots, r | \mathfrak{F}_n) \big| \mathfrak{F}_{n-1}\big) \\
&\quad = E\big(P(\xi_{\tau_1} \in B_1 | \mathfrak{F}_n) P(\xi_{\tau_i} \in B_i, i = 2, \ldots, r | \mathfrak{F}_n) \big| \mathfrak{F}_{n-1}\big)
\end{aligned}$$

$$= \frac{F(B_1 \cap A_n^{(1)})}{F(A_n^{(1)})} P(\xi_{\tau_i} \in B_i, i = 2, \dots, r | \mathfrak{F}_{n-1})$$
$$= \prod_{i=1}^{r} \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})}.$$

This completes the proof for ω-points in subsets of the form $\Omega'(s)$.

Establishing (18) *on* $\Omega''(s)$. Again for definiteness, we take s to be 1. For $\omega \in \Omega''(1)$, ξ_{τ_1} must take the value of the singleton $A_n^{(1)}$. Thus, by the induction assumption,

$$\begin{aligned} P(\xi_{\tau_i} \in B_i, i = 1, \dots, r | \mathfrak{F}_{n-1}) &= P(\xi_{\tau_1} \in B_1 | \mathfrak{F}_{n-1}) P(\xi_{\tau_i} \in B_i, i = 2, \dots, r | \mathfrak{F}_{n-1}) \\ &= P(\xi_{\tau_1} \in B_1 | \mathfrak{F}_{n-1}) \prod_{i=2}^{r} P(\xi_{\tau_i} \in B_i | \mathfrak{F}_{n-1}) \\ &= \prod_{i=1}^{r} \frac{F(B_i \cap A_n^{(i)})}{F(A_n^{(i)})}. \end{aligned}$$

This completes the proof for ω-points in subsets of the form $\Omega''(s)$, and completes the justification of the induction step.

The truncation step. To complete the proof, we must replace the given stopping times $\tau_1, \dots, \tau_r$ by a sequence of somewhat simpler disentangled indexical stopping times $\bar{\tau}_1(\nu), \dots, \bar{\tau}_r(\nu)$, $\nu = 1, 2, \dots$, which we shall write more compactly as $\bar{\tau}_1, \dots, \bar{\tau}_r$. Let

$$\bar{\tau}_1 = \begin{cases} \tau_1, & \text{if } \tau_1 < \nu, \\ \inf\{n \ge \nu : \xi_n \in A_\nu^{(1)}\}, & \text{if } \tau_1 \ge \nu, \end{cases}$$

and, recursively, for $i = 2, \dots, r$,

$$\bar{\tau}_i = \begin{cases} \tau_i, & \text{if } \tau_i < \nu, \\ \inf\{n \ge \nu : n > \max(\bar{\tau}_1, \dots, \bar{\tau}_{i-1}) \text{ and } \xi_n \in A_\nu^{(i)}\}, & \text{if } \tau_i \ge \nu. \end{cases}$$

Of course, $\xi_{\bar{\tau}_i} = \xi_{\tau_i}$ on $[\tau_i < \nu]$, so that

$$\begin{aligned} &\big|P(\xi_{\tau_i} \in B_i \text{ for } i = 1, \dots, r | \mathfrak{F}_{n-1}) - P(\xi_{\bar{\tau}_i} \in B_i \text{ for } i = 1, \dots, r | \mathfrak{F}_{n-1})\big| \\ &\quad \le P\big(\max(\tau_1, \dots, \tau_r) \ge \nu | \mathfrak{F}_{n-1}\big) \to 0 \qquad \text{as } \nu \to \infty. \end{aligned}$$

Likewise, $|P(\xi_{\tau_i} \in B_i | \mathfrak{F}_{n-1}) - P(\xi_{\bar{\tau}_i} \in B_i | \mathfrak{F}_{n-1})| \to 0$ as $\nu \to \infty$ for $i = 1, \dots, r$. Thus, if (18) holds for all n with τ_i's replaced by $\bar{\tau}_i$'s, then it holds without the replacement, and the proof is complete.

It is easy to see that $\bar{\tau}_1, \dots, \bar{\tau}_r$ are indexical: when $n \ge \nu$, $\bar{A}_n^{(i)} = A_\nu^{(i)} = \bar{A}_{n,1}^{(i)} + \bar{A}_{n,2}^{(i)}$, with one of the latter two an empty set, depending on n, and implicitly on ω, whatever is needed to enforce stopping at time $\bar{\tau}_i$. Likewise, it is easy to verify their pairwise disentanglement. Indeed, for $\bar{\tau}_i$ and $\bar{\tau}_j$ with $i < j$, for $n \geq \nu$, on $[\bar{\tau}_i \geq n,\ \bar{\tau}_j \geq n]$, we have $\bar{A}_n^{(j)} = \bar{A}_{n,2}^{(j)} = A_\nu^{(j)}$; that is, (8b) holds.

It remains to verify (18) when the τ_i's are replaced by $\bar{\tau}_i$'s. Because the induction step has already been established, it is enough to show this for each $n \geq \nu$, with the focus of attention on the ω-set $[\bar{\tau}_i \geq n$ for $i = 1, \dots, r]$. By the description of the $\bar{\tau}_i$'s, on $[\bar{\tau}_i \geq n,\ i = 1, \dots, r]$ (with $n \geq \nu$), $\bar{\tau}_1$ is the first time $k \geq n$ that $\xi_k \in A_\nu^{(1)}$, and for $i = 2, \dots, r$, $\bar{\tau}_i$ is the first time $k > \bar{\tau}_{i-1}$ that $\xi_k \in A_\nu^{(i)}$. In other words, given $\xi_1, \dots, \xi_{n-1}$ for which $\bar{\tau}_i \geq n$, $i = 1, \dots, r$, the differences $\bar{\tau}_1 - (n-1), \bar{\tau}_2 - \bar{\tau}_1, \dots, \bar{\tau}_r - \bar{\tau}_{r-1}$ are (conditionally) independent geometrically distributed random variables with respective parameters $F(A_\nu^{(i)})$, $i = 1, \dots, r$. It follows that

$$\begin{aligned} P(\xi_{\bar{\tau}_i} \in B_i,\ i = 1, \dots, r | \mathfrak{F}_{n-1}) &= \prod_{i=1}^r P(\xi_{\bar{\tau}_i} \in B_i | \mathfrak{F}_{n-1}) \\ &= \prod_{i=1}^r \frac{F(B_i \cap A_\nu^{(i)})}{F(A_\nu^{(i)})} \\ &= \prod_{i=1}^r \frac{F(B_i \cap \bar{A}_n^{(i)})}{F(\bar{A}_n^{(i)})}. \end{aligned}$$

This completes the proof. □

4. Converses. The intent in this section is to demonstrate that the indexical and disentanglement assumptions are natural. Given this limited objective, we shall state and prove Propositions 1 and 2 under conveniently restrictive assumptions. Proposition 1 shows that the stopping time τ must be indexical in order for the conditional distributions of ξ_τ to be as described in (7) for all $n \geq 1$ and all distributions F. Likewise, Proposition 2 shows that two indexical stopping times τ_1 and τ_2 must be disentangled in order for ξ_{τ_1} and ξ_{τ_2} to be conditionally independent under all distributions F.

PROPOSITION 1. *Assume $\mathfrak{X}$ is countable with $\mathfrak{B} = 2^{\mathfrak{X}}$ (the power set of $\mathfrak{X}$), and let τ be a stopping time which is a.s. finite for every distribution F that is fully supported on $\mathfrak{X}$. Further, suppose* (7) *holds for all such F, for every $B \in \mathfrak{B}$ and for $n = 1, 2, \dots$, where $A_1 = \mathfrak{X}$, and $A_n = A_n(\xi_1, \dots, \xi_{n-1})$ is a suitably chosen nonempty subset of $\mathfrak{X}$, defined on $[\tau \geq n]$ and depending on the observations $\xi_1, \dots, \xi_{n-1}$, for $n \geq 2$. Then the following hold*:

(i) *$A_{n+1} \subset A_n$ on $[\tau > n]$*;
(ii) *$[\tau \geq n$ and $\xi_n \notin A_n] \subset [\tau > n$ and $A_{n+1} = A_n]$*;

(iii) $[\tau \geq n \textit{ and } \xi_n \in A_n] = [\tau = n] + [\tau > n \textit{ and } \xi_n \in A_{n+1}(\xi_1, \ldots, \xi_n)]$;

(iv) *depending on* $\xi_1, \ldots, \xi_{n-1}$, *there exists a countable partition of* A_n *by subsets* $A_{n,i} = A_{n,i}(\xi_1, \ldots, \xi_{n-1})$, $i = 1, 2, \ldots$, *such that the following hold*:

(a) $[\tau \geq n \textit{ and } \xi_n \in A_{n,1}] = [\tau = n]$;

(b) $[\tau \geq n \textit{ and } \xi_n \in A_{n,i}] = [\tau > n \textit{ and } A_{n,i} = A_{n+1}]$ *for* $i \geq 2$.

PROOF. (i) Suppose, for given $\xi_1 = x_1, \ldots, \xi_{n-1} = x_{n-1}$, that $\tau > n$ when $\xi_n = x$ and, contrary to (i), that y belongs to $A_{n+1}(x_1, \ldots, x_{n-1}, x)$ but not to $A_n = A_n(x_1, \ldots, x_{n-1})$ for some y in $\mathfrak{X}$. Then, for any F supported on $\mathfrak{X}$, one obtains from (7) with $E := [\xi_1 = x_1, \ldots, \xi_{n-1} = x_{n-1}]$,

$$P(\xi_\tau = y|E) = F(\{y\}|A_n) = \frac{F(\{y\} \cap A_n)}{F(A_n)} = 0,$$

and the contradiction

$$\begin{aligned} P(\xi_\tau = y|E) &\geq P(\xi_\tau = y|E, \xi_n = x)F(\{x\}) \\ &= F\big(\{y\}|A_{n+1}(x_1, \ldots, x_{n-1}, x)\big)F(\{x\}) \\ &= \frac{F(\{y\})}{F(A_{n+1}(x_1, \ldots, x_{n-1}, x))}F(\{x\}) > 0. \end{aligned}$$

Thus $A_{n+1} \subset A_n$ on $[\tau > n]$, as asserted.

(ii) Suppose $\tau \geq n$ for given $\xi_1 = x_1, \ldots, \xi_{n-1} = x_{n-1}$ and, contrary to (ii), that $\tau = n$ when $\xi_n = x \notin A_n = A_n(x_1, \ldots, x_{n-1})$. Then by (7) with $E = [\xi_1 = x_1, \ldots, \xi_{n-1} = x_{n-1}]$,

$$0 = \frac{F(A_n^c \cap A_n)}{F(A_n)} = P(\xi_\tau \notin A_n|E) \geq P(\xi_n = x|E),$$

which leads to the contradiction $F(\{x_1\}) \cdots F(\{x_{n-1}\})F(\{x\}) = 0$, since no factor in the product can be zero when F is supported on $\mathfrak{X}$. Thus $\tau > n$ when $\tau \geq n$ and $\xi_n \notin A_n$. Next, suppose, contrary to the second part of (ii), that A_{n+1} is a proper subset of A_n when $\tau > n$ and $\xi_n = x \notin A_n = A_n(x_1, \ldots, x_{n-1})$. Let y belong to A_n but not to $A_{n+1} = A_{n+1}(x_1, \ldots, x_{n-1}, x)$ and choose for F a distribution which has $F(\{x\}) = 3/5$ and $F(\{y\}) = 1/5$, so that

$$P(\xi_\tau = y|E) = \frac{F(\{y\} \cap A_n)}{F(A_n)} \geq \frac{F(\{y\})}{1 - F(\{x\})} = \frac{1/5}{1 - 3/5} = \frac{1}{2}.$$

However, this is contradicted by

$$P(\xi_\tau = y|E) \leq P(\xi_n \neq x|E) = P(\xi_n \neq x) = 1 - F(\{x\}) = 2/5.$$

Finally, in view of (i), A_{n+1} must be a subset of A_n, and, hence, $A_{n+1} = A_n$. It follows that $\tau > n$ and $A_{n+1} = A_n$ when $\tau \geq n$ and $\xi_n \notin A_n$, as asserted.

(iii) Observe that

$$[\tau \geq n, \xi_n \in A_n] = [\tau = n, \xi_n \in A_n] + [\tau > n, \xi_n \in A_n].$$

By (ii), $[\tau = n,\ \xi_n \in A_n] = [\tau = n]$, and by (i), $[\tau > n,\ \xi_n \in A_n] \supset [\tau > n,\ \xi_n \in A_{n+1}]$. It remains to show $[\tau > n,\ \xi_n \in A_n] \subset [\tau > n,\ \xi_n \in A_{n+1}]$. For given $\xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}$ for which $\tau \geq n$, suppose to the contrary, $\xi_n = x \in A_n = A_n(x_1, \dots, x_{n-1})$ is such that $\tau > n$ and $x \notin A_{n+1}(x_1, \dots, x_{n-1}, x)$. Then consider a distribution F for which $F(\{x\}) = 2/3$, and observe, by (7), that

$$P(\xi_\tau = x|E) = \frac{F(\{x\} \cap A_n)}{F(A_n)} \geq F(\{x\}) = \frac{2}{3},$$

where $E = [\xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}]$. But this is contradicted by

$$\begin{aligned} P(\xi_\tau = x|E) &= P(\xi_\tau = x|E, \xi_n = x)F(\{x\}) \\ &\quad + P(\xi_\tau = x|E, \xi_n \neq x)\bigl(1 - F(\{x\})\bigr) \\ &\leq 0 \cdot (2/3) + 1 \cdot (1/3) = 1/3. \end{aligned}$$

Thus, $x \in A_{n+1}(x_1, \dots, x_{n-1}, x)$. That is, $\xi_n \in A_{n+1}$, as asserted, completing the proof of part (iii).

Part (a) *of* (iv). In view of part (iii), for given $\xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}$ for which $\tau \geq n$, it is apparent that the set $A_{n,1} = A_{n,1}(x_1, \dots, x_{n-1})$ described in part (a) of (iv) should be defined thusly:

$$A_{n,1} = \bigl\{x \in A_n : \tau = n \text{ when (together with the given } \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}) \ \xi_n = x\bigr\},$$

and nothing more needs to be said.

Part (b) *of* (iv). For given $\xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}$ for which $\tau \geq n$, the sets $A_{n,i} = A_{n,i}(x_1, \dots, x_{n-1})$, $i \geq 2$, which, together with $A_{n,1} = A_{n,1}(x_1, \dots, x_{n-1})$, describe the partitioning of $A_n = A_n(x_1, \dots, x_{n-1})$, are generated by the sets $A_{n+1}(x_1, \dots, x_{n-1}, x)$, $x \in A_n - A_{n,1}$. Since $\mathfrak{X}$ is countable, at most a countable number of sets can be generated in this way.

We note, as a consequence of (iii), that x is always a member of $A_{n+1}(x_1, \dots, x_{n-1}, x)$. Moreover, we claim for $x \in A_n - A_{n,1}$ that $A_{n+1}(x_1, \dots, x_{n-1}, x) \subset A_n - A_{n,1}$, which together with $x \in A_{n+1}(x_1, \dots, x_{n-1}, x)$ implies that the union of all of the sets $A_{n+1}(x_1, \dots, x_{n-1}, x)$ with $x \in A_n - A_{n,1}$ equals $A_n - A_{n,1}$. If, to the contrary of the claim, $y \in A_{n,1}(x_1, \dots, x_{n-1})$ for some $y \in A_{n+1}(x_1, \dots, x_{n-1}, x)$ (i.e., $\tau = n$ when $\xi_n = y$), then consider a distribution F for which $F(\{x\}) = F(\{y\}) = 2/5$, and observe with $E = [\xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}]$ that

$$P(\xi_\tau = y|E) = \frac{F(\{y\} \cap A_n)}{F(A_n)} \leq \frac{F(\{y\})}{F(\{x\}) + F(\{y\})} = \frac{1}{2}.$$

However, this is contradicted by

$$\begin{aligned} P(\xi_\tau = y|E) &\geq P(\xi_\tau = y|E, \xi_n = y)F(\{y\}) \\ &\quad + P(\xi_\tau = y|E, \xi_n = x)F(\{x\}) \\ &= 1 \cdot \left(\frac{2}{5}\right) + \frac{F(\{y\} \cap A_{n+1}(x_1, \dots, x_{n-1}, x))}{F(A_{n+1}(x_1, \dots, x_{n-1}, x))} \cdot \left(\frac{2}{5}\right) \\ &\geq \left(\frac{2}{5}\right) + \frac{2/5}{1} \cdot \left(\frac{2}{5}\right) = \frac{14}{25} > \frac{1}{2}. \end{aligned}$$

This proves the claim.

For the sets $A_{n+1}(x_1, \dots, x_{n-1}, x)$, $x \in A_n - A_{n,1}$, to be a partition of $A_n - A_{n,1}$, every pair of sets $A_{n+1}(x_1, \dots, x_{n-1}, x)$ and $A_{n+1}(x_1, \dots, x_{n-1}, y)$ $(x \neq y)$ must be either equal or disjoint. We will now show that the two sets are equal if, in addition, y is a member of the first set, and x is a member of the second. Also, we will show that the only other alternative is that neither of these memberships holds, and the two sets are disjoint. There are several cases to be ruled out:

Suppose $x \notin A_{n+1}(x_1, \dots, x_{n-1}, y)$, $y \notin A_{n+1}(x_1, \dots, x_{n-1}, x)$, but the intersection of these two sets contains a point $u \in A_n - A_{n,1}$. Clearly, $u \neq x$ and $u \neq y$. Consider a distribution F for which $F(\{x\}) = F(\{y\}) = F(\{u\}) = 3/10$. Then, by (7) with $E = [\xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}]$,

$$P(\xi_\tau = u|E) = \frac{F(\{u\} \cap A_n)}{F(A_n)} \leq \frac{F(\{u\})}{F(\{x\}) + F(\{y\}) + F(\{u\})} = \frac{1}{3}.$$

However, this is contradicted by

$$\begin{aligned} P(\xi_\tau = u|E) &\geq P(\xi_\tau = u|E, \xi_n = u)F(\{u\}) \\ &\quad + P(\xi_\tau = u|E, \xi_n = x)F(\{x\}) \\ &\quad + P(\xi_\tau = u|E, \xi_n = y)F(\{y\}) \\ &= \frac{F(\{u\})}{F(A_{n+1}(x_1, \dots, x_{n-1}, u))} F(\{u\}) \\ &\quad + \frac{F(\{u\})}{F(A_{n+1}(x_1, \dots, x_{n-1}, x))} F(\{x\}) \\ &\quad + \frac{F(\{u\})}{F(A_{n+1}(x_1, \dots, x_{n-1}, y))} F(\{y\}) \\ &\geq \frac{(3/10)^2}{1} + \frac{(3/10)^2}{1 - 3/10} + \frac{(3/10)^2}{1 - 3/10} = \frac{243}{700} > \frac{1}{3}. \end{aligned}$$

Another possibility is to suppose $x \notin A_{n+1}(x_1, \dots, x_{n-1}, y)$ but $y \in A_{n+1}(x_1, \dots, x_{n-1}, x)$, so that the intersection of these two sets contains the common point $y \in A_n - A_{n,1}$. Consider a distribution F for which $F(\{x\}) = F(\{y\}) = 5/11$. Then, by (7),

$$P(\xi_\tau = y|E) = \frac{F(\{y\} \cap A_n)}{F(A_n)} \le \frac{F(\{y\})}{F(\{x\}) + F(\{y\})} = \frac{1}{2}.$$

However, this is contradicted by

$$\begin{aligned} P(\xi_\tau = y|E) &\ge P(\xi_\tau = y|E, \xi_n = x)F(\{x\}) \\ &\quad + P(\xi_\tau = y|E, \xi_n = y)F(\{y\}) \\ &= \frac{F(\{y\})}{F(A_{n+1}(x_1, \dots, x_{n-1}, x))} F(\{x\}) \\ &\quad + \frac{F(\{y\})}{F(A_{n+1}(x_1, \dots, x_{n-1}, y))} F(\{y\}) \\ &\ge \frac{5/11}{1} \cdot \left(\frac{5}{11}\right) + \frac{5/11}{1 - 5/11} \cdot \left(\frac{5}{11}\right) = \frac{425}{726} > \frac{1}{2}. \end{aligned}$$

The case with the roles of x and y reversed is ruled out by a similar argument.

Thus, we have now shown that if $A_{n+1}(x_1, \dots, x_{n-1}, x)$ and $A_{n+1}(x_1, \dots, x_{n-1}, y)$ are not disjoint, then they must both contain x and y. Finally, suppose that $A_{n+1}(x_1, \dots, x_{n-1}, x)$ and $A_{n+1}(x_1, \dots, x_{n-1}, y)$ are neither disjoint nor equal, with both, necessarily, containing the points x and y. For definiteness, suppose the first set is not a subset of the latter and let $v \in A_n - A_{n,1}$ be a member of $A_{n+1}(x_1, \dots, x_{n-1}, x)$ that is not a member of $A_{n+1}(x_1, \dots, x_{n-1}, y)$. Clearly, $v \ne x$ and $v \ne y$. Also, by (iii), v is also a member of $A_{n+1}(x_1, \dots, x_{n-1}, v)$. Thus, $A_{n+1}(x_1, \dots, x_{n-1}, x)$ and $A_{n+1}(x_1, \dots, x_{n-1}, v)$ are not disjoint, and it follows (as we have shown in the context of x and y) that x is a member of $A_{n+1}(x_1, \dots, x_{n-1}, v)$. Consequently, x is a member of $A_{n+1}(x_1, \dots, x_{n-1}, v)$ and $A_{n+1}(x_1, \dots, x_{n-1}, y)$. Since the latter two sets are not disjoint, it follows that v is a member of $A_{n+1}(x_1, \dots, x_{n-1}, y)$, and this is a contradiction. This completes part (b) of (iv), and the proof of the proposition. □

PROPOSITION 2. *Assume $\mathfrak{X}$ is countable with $\mathfrak{B} = 2^{\mathfrak{X}}$ (the power set of $\mathfrak{X}$), and let τ_1 and τ_2 be indexical stopping times (with corresponding sets $A_n^{(i)}$ and $A_{n,k}^{(i)}$ for $k \ge 0$ and $i = 1, 2$), which are a.s. finite for every distribution F that is fully supported on $\mathbb{X}$. Suppose ξ_{τ_1} and ξ_{τ_2} are conditionally independent given $\mathfrak{F}_{n-1}$, for $n \ge 1$, as described in* (18) *with $r = 2$. Then τ_1 and τ_2 must be disentangled. That is, for each $n \ge 1$ and ω-point in $[\tau_1 \ge n,\ \tau_2 \ge n]$, there must be an index $s = s(\omega) \in \{1, 2\}$ which satisfies one of the three conditions* (8a)–(8c).

PROOF. Suppose, to the contrary, that there exists an $n \ge 1$ and values $x_1, \dots, x_{n-1}$ in $\mathfrak{X}$ such that $[\tau_1 \ge n]$ and $[\tau_2 \ge n]$ on $E := [\xi_1 = x_1, \dots, \xi_{n-1} =$

$x_{n-1}]$, with all of the following holding:

(24a) $A_n^{(1)} \not\subset A_{n,k}^{(2)}$ and $A_n^{(2)} \not\subset A_{n,k}^{(1)}$ for $k = 0, 2, 3, \dots$;

(24b) each of $A_n^{(1)}$ and $A_n^{(2)}$ includes either a (nonempty) stopping set or at least two continuation subsets (or both);

(24c) neither $A_n^{(1)}$ nor $A_n^{(2)}$ has cardinality 1.

By considering various choices for F, we will arrive at a contradiction. We shall only consider distributions F whose support is all of $\mathfrak{X}$.

We must systematically rule out a number of cases:

Case (i) $A_{n,1}^{(1)}$ *and* $A_{n,1}^{(2)}$ *are not disjoint.* If there exists a $u \in A_{n,1}^{(1)} \cap A_{n,1}^{(2)}$, then (by the conditional independence for $n+1$ in the first case of x_n below [cf. (18) and (3b)]),

$$P(\xi_{\tau_1} = u, \xi_{\tau_2} = u | E, \xi_n = x_n) = \begin{cases} \dfrac{F(\{u\})}{F(A_n^{(1)})} \dfrac{F(\{u\})}{F(A_n^{(2)})}, & \text{for } x_n \in A_{n,0}^{(1)} \cap A_{n,0}^{(2)}, \\ 1, & \text{for } x_n = u, \\ 0, & \text{otherwise.} \end{cases} \tag{25}$$

(The latter case of "otherwise" includes the subcase of stopping at time n by τ_i with $\xi_{\tau_i} = x_n \neq u$, $i = 1$ or 2; and it includes the subcase of $x_n \in A_{n,k}^{(i)}$ with $k \geq 2$, for $i = 1$ or 2, and, consequently, with $\tau_i \geq n+1$ and, necessarily, $\xi_{\tau_i} \in A_{n+1}^{(i)} = A_{n,k}^{(i)}$, thus ruling out $\xi_{\tau_i} = u$.) Hence,

$$P(\xi_{\tau_1} = u, \xi_{\tau_2} = u | E) = \frac{F(\{u\})}{F(A_n^{(1)})} \frac{F(\{u\})}{F(A_n^{(2)})} F(A_{n,0}^{(1)} \cap A_{n,0}^{(2)}) + F(\{u\}),$$

which, by conditional independence for n, must also equal

$$P(\xi_{\tau_1} = u | E) P(\xi_{\tau_2} = u | E) = \frac{F(\{u\})}{F(A_n^{(1)})} \frac{F(\{u\})}{F(A_n^{(2)})}.$$

So

$$F(A_n^{(1)}) F(A_n^{(2)}) = F(\{u\}) F(A_n^{(1)} \cup A_n^{(2)}). \tag{26}$$

On account of (24c) above, both $A_n^{(1)}$ and $A_n^{(2)}$ are strictly larger than the singleton set $\{u\}$. So one can easily find a distribution F which contradicts (26). For instance, one can make the factors on the left-hand side of (26) at least as large as $2F(\{u\}) > 0$ and, at the same time, the second factor on the right-hand side at most $3F(\{u\})$. This leads to the contradiction $2 \cdot 2 \leq 1 \cdot 3$, a contradiction that rules out case (i). So, at this point, we may add to assumptions (24a)–(24c) the following working assumption:

(24d) $A_{n,1}^{(1)} \cap A_{n,1}^{(2)} = \varnothing$.

Case (ii) *$A_{n,1}^{(1)}$ and $A_{n,0}^{(2)}$ are not disjoint, nor are $A_{n,1}^{(2)}$ and $A_{n,0}^{(1)}$.* If there exists a $u \in A_{n,1}^{(1)} \cap A_{n,0}^{(2)}$ and a $v \in A_{n,1}^{(2)} \cap A_{n,0}^{(1)}$, then

$$P(\xi_{\tau_1} = u,\, \xi_{\tau_2} = v | E,\, \xi_n = x_n) = \begin{cases} \dfrac{F(\{u\})}{F(A_n^{(1)})} \dfrac{F(\{v\})}{F(A_n^{(2)})}, & \text{for } x_n \in A_{n,0}^{(1)} \cap A_{n,0}^{(2)}, \\ \dfrac{F(\{v\})}{F(A_n^{(2)})}, & \text{for } x_n = u, \\ \dfrac{F(\{u\})}{F(A_n^{(1)})}, & \text{for } x_n = v, \\ 0, & \text{otherwise.} \end{cases}$$

[The arguments for the various subcases are similar to those used for (25).] Hence,

$$\begin{aligned} P(\xi_{\tau_1} = u,\, \xi_{\tau_2} = v | E) = {} & \frac{F(\{u\})}{F(A_n^{(1)})} \frac{F(\{v\})}{F(A_n^{(2)})} F\big(A_{n,0}^{(1)} \cap A_{n,0}^{(2)}\big) \\ & + \frac{F(\{v\})}{F(A_n^{(2)})} F(\{u\}) + \frac{F(\{u\})}{F(A_n^{(1)})} F(\{v\}), \end{aligned}$$

which simplifies to

$$P(\xi_{\tau_1} = u,\, \xi_{\tau_2} = v | E) = \frac{F(\{u\})}{F(A_n^{(1)})} \frac{F(\{v\})}{F(A_n^{(2)})} \big[1 + F\big(A_n^{(1)} \cap A_n^{(2)}\big)\big].$$

However, the assumption of conditional independence for n asserts that

$$P(\xi_{\tau_1} = u,\, \xi_{\tau_2} = v | E) = \frac{F(\{u\})}{F(A_n^{(1)})} \frac{F(\{v\})}{F(A_n^{(2)})}.$$

Thus $F(A_n^{(1)} \cap A_n^{(2)}) = 0$, which implies $A_n^{(1)} \cap A_n^{(2)} = \varnothing$. Since this conclusion is inconsistent with (24a), case (ii) is ruled out, and we may hereafter assume its converse as a working assumption. Because of the symmetry in case (ii), we shall, without loss of generality, add to our working assumptions:

(24e) $A_{n,1}^{(2)} \cap A_{n,0}^{(1)} = \varnothing$.

Case (iii) *$A_{n,1}^{(2)}$ is not empty.* From (24d) and (24e), we may conclude [see (6)] that $A_{n,1}^{(2)} \subset A_{n,2}^{(1)} + A_{n,3}^{(1)} + \cdots$. Hence, there exists a $u \in A_{n,1}^{(2)} \cap A_{n,k}^{(1)}$ for some $k \geq 2$, and

$$
\begin{aligned}
P(\xi_{\tau_1} = u, \xi_{\tau_2} = u | E, \xi_n = x_n) \\
= \begin{cases}
\dfrac{F(\{u\})}{F(A_n^{(1)})} \dfrac{F(\{u\})}{F(A_n^{(2)})}, & \text{for } x_n \in A_{n,0}^{(1)} \cap A_{n,0}^{(2)}, \\
\dfrac{F(\{u\})}{F(A_{n,k}^{(1)})}, & \text{for } x_n = u, \\
\dfrac{F(\{u\})}{F(A_{n,k}^{(1)})} \dfrac{F(\{u\})}{F(A_n^{(2)})}, & \text{for } x_n \in A_{n,k}^{(1)} \cap A_{n,0}^{(2)}, \\
0, & \text{otherwise.}
\end{cases}
\end{aligned}
$$

So

$$
\begin{aligned}
P(\xi_{\tau_1} = u, \xi_{\tau_2} = u | E) \\
= \frac{F(\{u\})}{F(A_n^{(1)})} \frac{F(\{u\})}{F(A_n^{(2)})} \\
\times \left[F(A_{n,0}^{(1)} \cap A_{n,0}^{(2)}) + \frac{F(A_n^{(1)}) F(A_n^{(2)})}{F(A_{n,k}^{(1)})} + \frac{F(A_n^{(1)}) F(A_{n,k}^{(1)} \cap A_{n,0}^{(2)})}{F(A_{n,k}^{(1)})} \right].
\end{aligned}
$$

Because of the assumed conditional independence for n, the terms within the bracket must add to unity, but, in fact, exceed unity when F is chosen to make $0 < F(\{u\}) \le F(A_{n,k}^{(1)}) < 1/9$, $F(A_n^{(1)} - A_{n,k}^{(1)}) > 1/3$ and $F(A_n^{(2)} - A_{n,k}^{(1)}) > 1/3$. [Note, $A_n^{(1)} - A_{n,k}^{(1)} \neq \varnothing$ by the first part of (24b), and $A_n^{(2)} - A_{n,k}^{(1)} \neq \varnothing$ by the second part of (24a).] So again, this case is contradicted, and we can strengthen working hypothesis (24e) to read:

(24f) $A_{n,1}^{(2)} = \varnothing$.

A consequence of (24f), together with (24b), is that $A_n^{(2)}$ includes two or more continuation subsets $A_{n,i}^{(2)}$ and $A_{n,j}^{(2)}$ with $i \neq j$, $i, j \ge 2$. Now suppose $u \in A_{n,k}^{(1)}$ is such that u belongs to the continuation subset $A_{n,i}^{(2)}$. [By (6), there must be such a k, $k \ge 0$.] Then, except when $k = 1$, the second condition in (24a) guarantees the existence of a second point $v \in A_n^{(2)}$ that is not a member of $A_{n,k}^{(1)}$. Clearly, the index j above can be chosen so that $v \in A_{n,j}^{(2)}$. (In the case $A_n^{(2)} - A_{n,i}^{(2)} \subset A_{n,k}^{(1)}$, we simply have the roles of i and j reversed.) Thus, in summary, there exist $u, v \in \mathfrak{X}$ such that $u \in A_{n,k}^{(1)} \cap A_{n,i}^{(2)}$, $v \in A_{n,l}^{(1)} \cap A_{n,j}^{(2)}$ with $i \neq j$, $i, j \ge 2$, and either $k = l = 1$ or $k \neq l$, $k, l \ge 0$. It is important to note that this summary statement *contains no conditional clauses*; once it is refuted (by appropriate demonstrations of contradictions), the proof of Proposition 2 is complete.

Because of symmetry, without loss of generality, we can assume that $k \geq l$. There are two cases to consider, each with two subcases: $k = 1$ and $k > 1$.

Case (iv) $k = 1$. Here, we have

$$P(\xi_{\tau_1} = u,\, \xi_{\tau_2} = u | E,\, \xi_n = x_n) = \frac{F(\{u\})}{F(A_{n,i}^{(2)})} \qquad \text{for } x_n = u,$$

so that

$$P(\xi_{\tau_1} = u,\, \xi_{\tau_2} = u | E) = \frac{F(\{u\})}{F(A_n^{(1)})} \frac{F(\{u\})}{F(A_n^{(2)})} \geq \frac{F(\{u\})}{F(A_{n,i}^{(2)})} F(\{u\}),$$

and, hence,

$$F(A_n^{(1)})\, F(A_n^{(2)}) \leq F(A_{n,i}^{(2)}). \tag{27}$$

For the subcase $l = 1$, (27) is contradicted by choosing F so that $F(\{v\}) > 1 - \varepsilon$, $F(\{u\}) < \varepsilon$, and $F(\{u\}) > \frac{1}{2} F(A_{n,i}^{(2)})$ for small $\varepsilon > 0$.

For the subcase $l = 0$, (27) is contradicted by choosing F as above, but with the restriction on v replaced by the restrictions $F(\{v\}) > \frac{1}{2} - \varepsilon$ and $F(\{y\}) > \frac{1}{2} - \varepsilon$, where $y \in A_n^{(1)} - A_{n,i}^{(2)}$. Such a y must exist on account of the first part of (24a), and we note that $y \neq u$ and $y \neq v$.

Case (v) $k \geq 2$. Note that

$$P(\xi_{\tau_1} = u,\, \xi_{\tau_2} = u | E,\, \xi_n = x_n) = \frac{F(\{u\})}{F(A_{n,k}^{(1)})} \frac{F(\{u\})}{F(A_{n,i}^{(2)})} \qquad \text{for } x_n = u,$$

so that

$$P(\xi_{\tau_1} = u,\, \xi_{\tau_2} = u | E) = \frac{F(\{u\})}{F(A_n^{(1)})} \frac{F(\{u\})}{F(A_n^{(2)})} \geq \frac{F(\{u\})}{F(A_{n,k}^{(1)})} \frac{F(\{u\})}{F(A_{n,i}^{(2)})} F(\{u\}),$$

and, hence,

$$F(A_n^{(1)})\, F(A_n^{(2)})\, F(\{u\}) \leq F(A_{n,k}^{(1)})\, F(A_{n,i}^{(2)}). \tag{28}$$

For the subcase $l \geq 1$, we have $u, v \in A_n^{(1)} \cap A_n^{(2)}$, and (28) is contradicted by choosing F so that $F(\{v\}) > 1 - \varepsilon$, $F(\{u\}) < \varepsilon$, $F(\{u\}) > \frac{1}{2} F(A_{n,k}^{(1)})$ and $F(\{u\}) > \frac{1}{2} F(A_{n,i}^{(2)})$ for small $\varepsilon > 0$.

For the subcase $l = 0$, there are two possibilities:

(a) If $A_{n,k}^{(1)} \subset A_{n,i}^{(2)}$, the first condition in (24a) guarantees the existence of a $y \in A_n^{(1)} - A_{n,i}^{(2)}$, necessarily with $y \notin A_{n,k}^{(1)}$, $y \neq u$ and $y \neq v$ (since $v \notin A_n^{(1)}$). Then (28) is contradicted by choosing F so that $F(\{v\}) > \frac{1}{2} - \varepsilon$, $F(\{y\}) > \frac{1}{2} - \varepsilon$, $F(\{u\}) < \varepsilon$, $F(\{u\}) > \frac{1}{2} F(A_{n,k}^{(1)})$ and $F(\{u\}) > \frac{1}{2} F(A_{n,i}^{(2)})$ for small $\varepsilon > 0$.

(b) If $A_{n,k}^{(1)} \not\subset A_{n,i}^{(2)}$, let $u' \in A_{n,k}^{(1)} - A_{n,i}^{(2)}$. By the first part of (24b), $A_{n,m}^{(1)} \neq \varnothing$ for some index m, $0 < m \neq k$.

If $A_{n,m}^{(1)} \not\subset A_{n,i}^{(2)}$, choose $y \in A_{n,m}^{(1)} - A_{n,i}^{(2)}$. Then (28) is contradicted by choosing F as in part (a) above.

Alternatively, if $A_{n,m}^{(1)} \subset A_{n,i}^{(2)}$, choose $y \in A_{n,m}^{(1)}$. The case of $m = 1$ can be treated as for $k = 1$ in Case (iv) above (with u replaced by y). For $m > 1$, one obtains the inequality

$$F(A_n^{(1)})\,F(A_n^{(2)})\,F(\{y\}) \le F(A_{n,m}^{(1)})\,F(A_{n,i}^{(2)}) \tag{29}$$

in the same manner (28) is obtained, but with m replacing k, and y replacing u. Finally, (29) is contradicted by choosing F so that $F(\{v\}) > \frac{1}{2} - \varepsilon$, $F(\{u'\}) > \frac{1}{2} - \varepsilon$, $F(\{y\}) < \varepsilon$, $F(\{y\}) > \frac{1}{2}F(A_{n,m}^{(1)})$ and $F(\{y\}) > \frac{1}{2}F(A_{n,i}^{(2)})$ for small $\varepsilon > 0$. $\square$

APPENDIX

The task here is to verify that the ω-sets described in (19), (20) and (21) are "universally measurable sets" in that they are measurable under the completion of $\mathfrak{F}_{n-1}$ for *every* distribution F. To accomplish this, we first need to establish some general mathematical facts concerning the product measurable space arising from two Polish spaces $(\mathfrak{Y}, \mathfrak{C})$ and $(\mathfrak{Z}, \mathfrak{D})$. If $E \subset \mathfrak{Y} \times \mathfrak{Z}$, is a Borel set (a set in the product σ-field $\mathfrak{C} \times \mathfrak{D}$), then the set

$$E' = \{y \in \mathfrak{Y} : (y, z) \in E \text{ for some } z \in \mathfrak{Z}\}$$

is called the *projection* of E on $\mathfrak{Y}$. Following an error by Lebesgue in 1905, Souslin demonstrated about 10 years later that a projection set need not be measurable. Later, this observation led to the defining and general study of *analytic sets*, of which E' is an example. From this, it was discovered that analytic sets are *universally measurable sets* in the sense that they are measurable in the completion of $\mathfrak{C}$ under every distribution on $(\mathfrak{Y}, \mathfrak{C})$. The reader is referred to Bruckner, Bruckner and Thomson (1997) for a readable discussion of these matters.

Now let Γ denote the set of cardinals and, for $M \subset \Gamma$, let

$$E_{[M]} := \{y \in \mathfrak{Y} : \mathrm{card}(E_y) \in M\},$$

where E_y denotes the y-section of E. Further, let $M_n := \{\gamma \in \Gamma : \gamma \ge n\}$, $n = 0, 1, 2, \dots$, and observe that $E' = E_{[M_1]}$; and, of course, $E_{[M_0]} = \mathfrak{Y}$. Finally, let $\mathfrak{M}$ denote the smallest σ-field containing the sets M_n, $n \ge 0$.

PROPOSITION 3. *Every subset of $\mathfrak{Y}$ of the form $E_{[M]}$, $E \in \mathfrak{C} \times \mathfrak{D}$, $M \in \mathfrak{M}$, is a universally measurable set.*

PROOF. It is easily verified that it is enough to show that sets of the form $E_{[M_n]}$ are universally measurable sets since $\mathfrak{M}$ consists of those sets M of countable cardinality and those whose complement is of countable cardinality.

We have already dealt with $E_{[M_0]}$ and $E_{[M_1]}$. We shall demonstrate that $E_{[M_2]}$ is a universally measurable set; the remaining sets $E_{[M_n]}$, $n \geq 3$, can be dealt with similarly. To this end, introduce the Polish space $(\mathfrak{Z}', \mathfrak{D}') := (\mathfrak{Z} \times \mathfrak{Y} \times \mathfrak{Z}, \mathfrak{D} \times \mathfrak{C} \times \mathfrak{D})$ and observe for the Borel set $\hat{E} \subset \mathfrak{Y} \times \mathfrak{Z}'$ of the form $\{(y_1, z_1, y_2, z_2) \in E \times E : y_1 = y_2 \text{ and } z_1 \neq z_2\}$ that $E_{[M_2]} = \hat{E}_{[M_1]}$. Thus, $E_{[M_2]}$ is an analytic subset of $\mathfrak{Y}$ and, hence, is a universally measurable set. $\square$

Finally, we return our attention to the ω-sets described in (19), (20) and (21). Each of these sets, as does T in (19)–(21), depends on $\omega = (\omega_1, \omega_2, \ldots)$ through its first $n-1$ components only. So, for the purpose of this discussion, we may (and will) view them as a subset of $\mathfrak{Y} = \mathfrak{X}^{(n-1)}$, the $(n-1)$-fold product of $\mathfrak{X}$. Further, let $\mathfrak{Z} = \mathfrak{X}$. Here, $\mathfrak{C} = \mathfrak{B}^{(n-1)}$ and $\mathfrak{D} = \mathfrak{B}$, where $\mathfrak{B}$ is the σ-field attached to $\mathfrak{X}$. For each distribution F on $(\mathfrak{X}, \mathfrak{B})$, there is a corresponding distribution on $(\mathfrak{Y}, \mathfrak{C})$. Together, these give rise to a collection $\bar{\mathfrak{C}}$ (a σ-field) of universally measurable sets. We now argue that the sets in (19), (20) and (21) are members of $\bar{\mathfrak{C}}$. Since $\Omega(s,k) + (\mathfrak{Y} - T) = E_{[\varnothing]}$ when $E = \{(\omega_1, \ldots, \omega_{n-1}, x) : x \in A_n^{(s)} \text{ and } x \notin A_{n,k}^{(3-s)}\}$, it follows that $\Omega(s,k) = E_{[\varnothing]} \cap T \in \bar{\mathfrak{C}}$. Likewise, $\Omega'(s) = E_{[\varnothing]} \cap T \in \bar{\mathfrak{C}}$ when $E = \{(\omega_1, \ldots, \omega_{n-1}, x) : x \in A_n^{(s)} \text{ and } x \notin A_{n,2}^{(s)}\}$. Finally, $\Omega''(s) = E_{[\{1\}]} \cap T \in \bar{\mathfrak{C}}$ when $E = \{(\omega_1, \ldots, \omega_{n-1}, x) : x \in A_n^{(s)}\}$. So, to conclude, all of the sets in (19), (20) and (21) are universally measurable sets, as required.

Acknowledgment. We are indebted to S. R. S. Varadhan for a helpful suggestion that led us to the formulation and proof of Proposition 3.

REFERENCES

BRUCKNER, A. M., BRUCKNER, J. B. and THOMSON, B. S. (1997). *Real Analysis.* Prentice-Hall, Upper Saddle River, NJ.

DEHEUVELS, P. (1983). The strong approximation of extremal processes II. *Z. Wahrsch. Verw. Gebiete* **62** 7–15.

DOOB, J. L. (1936). Note on probability. *Ann. of Math.* **37** 363–367.

ENGELEN, R., TOMMASSEN, P. and VERVAAT, W. (1988). Ignatov's theorem: a new and short proof. *J. Appl. Probab.* **25A** 229–236.

GOLDIE, C. M. and ROGERS, L. C. G. (1984). The k-record processes are i.i.d. *Z. Wahrsch. Verw. Gebiete* **67** 197–211.

IGNATOV, Z. (1977). Ein von der Variationsreihe erzeuger Poissonscher Punctprozess. *Annuaire Univ. Sofia Fac. Math. Méch.* **71** 79–94. (Published 1986.)

RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

ROGERS, L. C. G. (1989). Ignatov's theorem: an abbreviation of the proof of Engelen, Tommassen and Vervaat. *Adv. Appl. Probab.* **21** 933–934.

SAMUELS, S. M. (1992). An all-at-once proof of Ignatov's theorem. *Contemp. Math.* **125** 231–237.

STAM, A. J. (1985). Independent Poisson processes generated by record values and interrecord times. *Stochastic Process. Appl.* **19** 315–325.

YAO, Y.-C. (1997). On independence of k-record processes: Ignatov's theorem revisited. *Ann. Appl. Probab.* **7** 815–821.

G. SIMONS
DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27599-3260
E-MAIL: simons@stat.unc.edu

Y.-C. YAO
INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI
TAIWAN
E-MAIL: yao@stat.sinica.edu.tw

L. YANG
DEPARTMENT OF STATISTICS
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824
E-MAIL: yang@stt.msu.edu