The Annals of Statistics
1997, Vol. 25, No. 3, 1284–1310

AVERAGE RUN LENGTH TO FALSE ALARM FOR SURVEILLANCE SCHEMES DESIGNED WITH PARTIALLY SPECIFIED PRE-CHANGE DISTRIBUTION

BY LOUIS GORDON[1,2] AND MOSHE POLLAK[2]

Filoli Information Systems and Hebrew University of Jerusalem

Observations are taken independently and sequentially. Detection of a change in distribution is studied when the problem has an invariance structure. The prechange distribution is assumed to be a member of a specified family and is assumed known up to a nuisance parameter. We provide a general method of constructing surveillance schemes in the presence of a nuisance parameter and give sufficient conditions for approximating their average run lengths to false alarm. Applications include detecting a change in scale of i.i.d. gamma variates with unknown initial scale, detecting a change in location of i.i.d. normal variates with unknown initial mean, and a non-parametric scheme based on ranks for detecting a change to a stochastically larger distribution.

1. Introduction. The problem of sequentially detecting a changepoint has been of interest in the context of process control for over half a century. Shewhart (1931) introduced the notion of control charts with associated "3σ" limits for the problem of detecting a large and sudden shift in mean in a set of independent normally distributed observations with known variance and known initial mean. Cusum control charts, introduced by Page (1954) provide a substantial increase in sensitivity of detection over Shewhart charts. Both procedures are in widespread industrial use today. Our mathematical understanding of cusum procedures is profoundly advanced in the paper of Lorden (1971).

We consider observations $X_1, X_2, \ldots$ taken sequentially. Under the measure P_∞, the observations are i.i.d. with continuous distribution F_0. Under the measure P_k, prechange observations $X_1, \ldots, X_{k-1}$ are i.i.d. with distribution F_0, independent of postchange observations $X_k, \ldots$; the latter are i.i.d. with continuous distribution F_1. At the nth observation, the cusum procedure requires us to compute $\max_{k\le n} dP_k/dP_\infty(X_1, \ldots, X_n)$ and to assert a change has occurred when the statistic first exceeds some critical level.

We desire a procedure with associated stopping rule which under P_k will stop with high probability shortly after the kth observation—ensuring rapid

Received September 1990; revised September 1996.
[1]Research supported in part by the John Simon Guggenheim Memorial Foundation and U.S.–Israel Binational Science Foundation.
[2]Research supported in part by NSF Grant DMS-90-05833 and U.S.–Israel Binational Science Foundation Grant 89-00115.
AMS 1991 *subject classifications*. Primary 62L10; secondary 62N10.
Key words and phrases. Cusum, change-point, disruption, invariance.

detection of a change. It is of course undesirable to stop before time k under the P_k measure—a false alarm. Naturally, we wish to control the tendency to issue false alarms, measured by the average run length (ARL) to false alarm.

When F_0 and F_1 are completely known to the statistician, increasingly strong optimality properties of the cusum procedure have been shown by Lorden (1971), Moustakides (1986) and Ritov (1990).

A competitor to the cusum technique is due to Shiryayev (1963) and Roberts (1966) when prechange and postchange distributions are completely specified. They both suggest computing the sum of likelihood ratios $R_n = \sum_{k=1}^{n} d P_k / d P_\infty(X_1, \dots, X_n)$ at time n and asserting that a change has already occurred when R_n first exceeds a specified level A. Pollak (1985) shows that the Shiryayev–Roberts procedures have strong asymptotic optimality properties. Also, see Yakir (1997).

If the postchange distribution is unknown, classical theory essentially calls for choosing a representative distribution F_1 and employing a control chart based on F_0 and the representative F_1. Another approach is to employ mixtures; see Pollak (1987).

Much less is known when the prechange distribution is not completely specified; for example, if the form of the prechange distribution is specified up to an unknown nuisance parameter. By analogy with hypothesis testing, a natural approach is to reduce the compound decision problem by using invariant statistics. We desire a detection procedure invariant under some set of transformations $x \to g(x)$. Given a set of invariant statistics $T_n = T_n(X_1, \dots, X_n)$, we can compute the sum of invariant likelihood ratios $R_n = \sum_{k=1}^{n} d P_k / d P_\infty(T_1, \dots, T_n)$ determined by the joint densities of the invariant statistics at time n, and assert that a change has already occurred when R_n first exceeds a specified level A. For technical reasons, this Shiryayev–Roberts statistic is more amenable to analysis than the analogous cusum statistic.

For example, Pollak and Siegmund (1991) studied a Shiryayev–Roberts procedure invariant under the translation group and so treated the problem of detecting a unit shift in mean when prechange observations are i.i.d. $\mathcal{N}(\mu, 1)$ with μ unknown and postchange observations are i.i.d. $\mathcal{N}(\mu + \delta, 1)$.

We here study the ARL to false alarm of such invariant procedures. While conceptually straightforward, the technical difficulties involved are formidable. Because one loses independence, the likelihood ratio of the invariant statistics is no longer a product of univariate ratios, so one must find a computationally useful form for the likelihood ratio. Because one loses identical distribution, the tools of renewal theory are not directly applicable. While these difficulties can be overcome on a case-by-case basis, a substantial number of similarities are common to many cases. It therefore is useful to present a general formulation which isolates the difficulties unique to specific problems. This we do with Theorem 1 stated in Section 2 and proved in Section 4. Examples are given in Section 3.

Our first example involves detection of a sudden change in scale in a sequence of independent gamma variates with known shape and unknown initial scale. This problem is invariant under the group of scale-change transformations. The nuisance parameter is the initial scale. Such problems include the detection of a halving in the mean of exponential variates from unknown baseline, or the doubling in variance of normally distributed observations having the same known mean prechange and postchange. The former problem was studied by Lorden and Eisenberger (1973). For the latter problem in a practical context, see Wilson, Griffiths, Kemp, Nix and Rowlands (1979).

Our second example involves detection of a shift in normal mean from unknown baseline. The problem is invariant under the group of shifts. The nuisance parameter is the prechange mean. Both prechange and postchange behavior of the invariant Shiryayev–Roberts statistic is given in Pollak and Siegmund (1991). We briefly indicate how results there imply the hypotheses of Theorem 1.

Our third example involves the detection of a stochastically larger distribution following disruption. The problem is invariant under the group of strictly increasing transformations. The nuisance parameter is the pre-change distribution. We present without proof an invariant likelihood ratio based on ranks and the asymptotic ARL to false alarm. The proof, using Theorem 1, and an analysis of the postchange ARL to detection is deferred to a companion paper, Gordon and Pollak (1995). Another application of Theorem 1 appears in Gordon and Pollak (1994). Studied there are sequential nonparametric procedures in which the prechange distribution is symmetric, with known center of symmetry.

2. The schemes and their ARL to false alarm. We consider two families of probability density functions. The first, $\mathscr{H}_0 = \{f_0(\cdot|\eta) \,|\, \eta \in H_0\}$ is assumed entirely indexed by a nuisance parameter η. The second family $\mathscr{H} = \{f(\cdot|\eta, \alpha) \,|\, (\eta, \alpha) \in H_0 \times A\}$ is assumed to be indexed by both the same nuisance parameter η as specifies $\mathscr{H}_0$ and possibly by another parameter α.

Let $X_1, X_2, \dots$ be independent observations under any of an indexed set of probability measures $P_1^{\eta, \alpha}, P_2^{\eta, \alpha}, \dots, P_\infty^{\eta} = P_\infty^{\eta, \alpha}$. Under P_∞^{η}, they are identically distributed with marginal density $f_0(\cdot|\eta)$, where the nuisance parameter η is unknown. Under $P_\nu^{\eta, \alpha}$, observations X_j for $j < \nu$ are i.i.d., again with marginal density $f_0(\cdot|\eta)$, while the subsequent observations $X_\nu, X_{\nu+1}, \dots$ are i.i.d. with marginal density $f(\cdot|\eta, \alpha)$.

We assume there exists a function $\alpha(\eta)$ such that when X has density $f_0(\cdot|\eta)$ then the likelihood ratio random variable $f(X|\eta, \alpha(\eta))/f_0(X|\eta)$ has distribution that does not depend on η. Note that the nuisance parameter η may appear in the likelihood ratio random variable even though its distribution does not depend on η. While this assumption seems restrictive, it is satisfied in the three examples described at the end of Section 1.

Assume also that there exists a sequence of nontrivial invariant statistics $\{T_n\}$ which are possibly vector-valued, where $T_n = T_n(X_1, \dots, X_n)$ are such that the distribution of $\{T_n\}$ induced by any $P_\nu^{\eta, \alpha(\eta)}$ does not depend on η. We write $\mathscr{F}_n$ for the σ-algebra generated by the statistics $\{T_1, \dots, T_n\}$. As usual, $\mathscr{F}_0$ is the trivial σ-algebra. Because of the assumed invariance, so long as we only study properties determined by the sequence $T_1, T_2, \dots$ we may specify in our computations some canonical nuisance parameter η_0 whose choice is solely a matter of convenience. We write P_ν for $P_\nu^{\eta_0, \alpha(\eta_0)}$ and E_ν for the expectation corresponding to P_ν.

In this paper, we choose a distinguished family of densities $f(\cdot|\eta, \alpha(\eta))$ to represent the postchange distribution. That is, we concentrate on the leading special case for which in the presence of an unknown nuisance parameter one wishes to detect a change from observations generated with density $f_0(\cdot|\eta)$ prechange to those generated with density $f(\cdot|\eta, \alpha(\eta))$ postchange.

The likelihood ratio $\Lambda_k^n(T_1, \dots, T_n) = dP_k/dP_\infty(T_1, \dots, T_n)$ is well defined and has a distribution under P_∞ that does not depend on the unknown parameter η. To implement a Shiryayev–Roberts procedure we define $R_n = \sum_{k=1}^n \Lambda_k^n$ and the stopping rule $N_A = \min\{n|R_n \geq A\}$. Without loss, we may choose a convenient η_0 and do all calculations as if we knew that $\eta = \eta_0$ were truly to be the nuisance parameter which specifies both prechange and postchange distributions. Write $f_1(\cdot) = f(\cdot|\eta_0, \alpha(\eta_0))$ and $f_0(\cdot) = f_0(\cdot|\eta_0)$. If prechange and postchange densities are completely specified, Pollak (1987) shows that, given level A, the ARL to false alarm is asymptotically $A\Delta(f_0(\cdot|\eta), f(\cdot|\eta, \alpha(\eta)))$, where $\Delta(f(\cdot|\eta), f(\cdot|\eta, \alpha(\eta))$ will be described below. By assumption, $\Delta(f_0(\cdot|\eta), f(\cdot|\eta, \alpha(\eta)))$ does not depend on η, so that we may denote this quantity as $\Delta(f_0, f_1)$.

Specifically, we obtain $\Delta(f_0, f_1)$ as follows. Let $Z_i = \log(f_1(X_i)/f_0(X_i))$ with associated hitting time $M_b = \inf\{n \mid \sum_{i=1}^n Z_i \geq b\}$, where $M_b = \infty$ if no such n exists. Recall that E_1 denotes expectation when the X_i are i.i.d. with density f_1. From Pollak (1987),

$$\Delta(f_0, f_1) = 1 \Big/ \lim_{b\to\infty} E_1\left\{\exp\left(-\left[\sum_{i=1}^{M_b} Z_i\right] + b\right)\right\},$$

which is computable by renewal theoretic methods. See, for example, Siegmund (1985), Chapter VIII.

Our theorem shows that—under suitable conditions—the invariant Shiryayev–Roberts statistic has the same limiting constant of proportionality $\Delta(f_0, f_1)$. Therefore, asymptotically, it has the same ARL to false alarm as the Shiryayev–Roberts procedure with nuisance parameter known and specified.

The key observation is that $R_n - n$ is a zero-expectation martingale under the prechange measure P_∞. Therefore, given critical level A, we study the expected time to false alarm $E_\infty\{N_A\}$ by studying the expected level $E_\infty\{R_{N_A}\}$ at which a false alarm is asserted. We have found that a common approach can be applied to obtain the limit of $E_\infty\{R_{N_A}\}/A$ as $A \to \infty$. We prove below that this limit equals $\Delta(f_0, f_1)$ whenever we can verify the three conditions formally given in the statement of Theorem 1.

Condition 1A requires that there be little contribution from likelihood ratios whose putative changepoint k is far removed from the ends of the sequence of observations. This would be clear in the i.i.d. case where the log-likelihood is a random walk with negative drift. The hypothesis thus says that there should be enough information about the original data in the invariant statistics $T_1, \dots, T_n$ to make the invariant likelihood behave similarly to the fully specified likelihood. The restriction to values k of intermediate magnitude comes from situations like the normal shift problem in which the measures P_∞ and P_1 are identical.

Condition 1B has the heuristic interpretation that, given η_0, when lots of data are observed prechange, the invariant statistics $T_1, \dots, T_n$ permit a near-perfect reconstruction of the data $X_1, \dots, X_n$. Hence, conditional on the past, the next increment in invariant log-likelihood is almost the fully specified log-likelihood. In symbols, $R_{n+1} \approx (1 + R_n) f_1(X_{n+1})/f_0(X_{n+1})$. When n and R_n are both large we should therefore expect that for a short while the statistics R_{n+j} should be close to $R_n \prod_{i=1}^{j} [f_1(X_{n+i})/f_0(X_{n+i})]$, enabling us to approximate the excess over the boundary—if it should occur rapidly—by that appropriate from renewal theory for the i.i.d. fully specified log-likelihood random variables.

Conditions 1A and 1B are used only once during the proof, to establish Lemma 6 of Section 4. In contrast, Condition 1C is used many times. As with the other conditions, Condition 1C also requires the invariant log-likelihood ratio to behave qualitatively like the log-likelihood ratio in the fully specified i.i.d. case. Specifically, relative to A, the overshoot should have bounded expectation and should be uniformly integrable. Although the other two conditions are permitted to fail with small probability, Condition 1C must hold with probability 1. This allows us to control the contribution of the conditional expectation on exceptional sets. While we often use Condition 1C to bound expectations, uniform integrability is used only to prove Lemma 10, in the argument following (17).

THEOREM 1. *Suppose that the following three conditions hold.*

1A. *Let* $0 < \varepsilon_1, \varepsilon_2 < 1$ *be given. There then exist positive constants* a_1, a_2 *and* a_3 *depending on* ε_1 *and* ε_2 *such that for all* $n \geq 1$,

$$P_\infty \left\{ \sup_{n\varepsilon_1 \leq k \leq n(1-\varepsilon_2)} \Lambda_k^n > e^{-a_1 n} \right\} < a_2 e^{-a_3 n}.$$

1B. *Let* $0 < \varepsilon < 1$ *be given. There then exist positive constants* $\theta < 1, b_1, b_2$ *and a set* B_ε, *all depending only on* ε *such that for all* $n \geq 1$,

$$P_\infty \left\{ X_{n+1} \in B_\varepsilon \text{ and } \max_{(1-\theta)n \leq k \leq n+1} \left| 1 - \frac{\Lambda_k^{n+1}}{\Lambda_k^n} \Big/ \frac{f_1(X_{n+1})}{f_0(X_{n+1})} \right| > \varepsilon \right\} \leq b_1 e^{-b_2 n}$$

and $P_\infty \{X_{n+1} \notin B_\varepsilon\} < \varepsilon$.

1C. *For* $t \geq 1$ *there exist finite functions* $A_0(t)$ *and* $\kappa(t)$ *such that* $\kappa(t) \to 0$ *as* $t \to \infty$, *and such that if* $M \geq 0$ *is any stopping time adapted to* $\mathscr{F}_n$ *and* $N_A^M = N = \min\{n | \sum_{k=M+1}^n \Lambda_k^n \geq A\}$ *then*

$$E_\infty \left\{ \left[\sum_{k=M+1}^{N} \Lambda_k^N \right] I_{\{\sum_{k=M+1}^N \Lambda_k^N \geq At\}} \,\middle|\, M; T_1, \dots, T_M \right\} \leq \kappa(t) A$$

uniformly for all $A > A_0(t)$, *all* $t \geq 1$ *and all stopping times* M.

Suppose also that the log-likelihood ratio $\log(f_1(X)/f_0(X))$ *has a continuous distribution when the prechange density is* $f_0(\cdot)$. *We may then conclude that* $\lim_{A\to\infty} E_\infty\{N_A/A\} = \Delta(f_0, f_1)$.

An immediate consequence of Condition 1C is that $\kappa(1) \geq 1$. An alternative condition which implies Condition 1C is that for some δ positive, $A^{-(1+\delta)}E_\infty\{(\sum_{k=M+1}^{N}\Lambda_k^N)^{1+\delta} \mid M, T_1, \ldots, T_M\} < \kappa'$ for all $A > A_0$. To see this, let X be any random variable. By Holder's inequality, $E\{XI_{\{X>t\}}\} < (E\{X^{1+\delta}\})^{1/(1+\delta)}(P\{X > t\})^{\delta/(1+\delta)}$. By Markov's inequality, the latter upper bound is in turn bounded by $t^{-\delta}E\{X^{1+\delta}\}$.

A more arduous approach would be to treat mixtures across orbits of the set of possible postchange distributions. While such elaboration appears feasible, it would lengthen an already long argument.

3. Examples.

3.1. *Detecting a change in a gamma sequence with unknown initial scale.* Our first application is to detecting a change in scale of gamma-distributed observations with known shape parameter. The results are formally presented in Theorem 2. The theorem covers exponentially distributed observations, as would occur if one were monitoring a Poisson process with unknown mean interarrival time for a sudden decrease in mean. The theorem also pertains to the chi-square distribution with a single degree of freedom, as is the case if one were monitoring normally distributed variates with known mean and unknown initial variance for a sudden increase in variance. We believe both problems to be of practical interest. The problem we treat is invariant under the group of scale-change transformations. The nuisance parameter is the initial unknown scale of the gamma-distributed observations.

The problem of detecting a sudden change in exponential mean is also treated in Lorden and Eisenberger (1973), where the statistics T_j of Theorem 2 are also used. The value of using invariant Shiryayev–Roberts statistics lies in both the ease with which an appropriate procedure is obtained from a general viewpoint, and in the applicability of our Theorem 1.

A number of technical points are illustrated by this example. First, P_∞ and P_1 induce identical distributions for the invariant statistics. Hence the likelihood ratio Λ_1^n equals 1 identically for all n. This is why Condition 1A allows us to neglect Λ_k^n only for k which are neither too large nor too small.

Second, Condition 1C is trivially verified in case the postchange distribution involves postchange scale parameter $\alpha > 1$, implying a decrease in mean, because the fully specified likelihood ratio is bounded. However, the verification is nontrivial in case $\alpha < 1$. The style of argument used in the latter case is also of interest. We have found it highly effective to argue the monotonicity of the Shiryayev–Roberts statistic relative to the last observation, and then to argue conditionally with respect to the threshold established by monotonicity. Similar arguments are used in Pollak and Siegmund (1991).

Third, although this paper deals only with ARL to false alarm, the procedures derived from invariance considerations often have desirable operating characteristics postchange as well. For example, it is possible to prove for the gamma scale-change problem that, for changes occurring even relatively early, the procedures of Theorem 2 have very strong conditional upper bounds for expected time to detection postchange. These bounds are comparable to Lorden's (1971) minimax bounds for cusum procedures. (The key observation is that if the first postchange observation is seen at time k, then the prechange statistics R_j for $j < k$ are independent of the sum $\sum_{j=1}^{k-1} X_j$.) A similar result due to Pollak and Siegmund (1991) holds in the context of our next example, presented later.

THEOREM 2. *Let $F(\cdot)$ be the gamma distribution function with shape β and scale 1, having density $f(x) = \Gamma^{-1}(\beta)x^{\beta-1}e^{-x}$.*

(i) *The likelihood ratio of the sequence of maximal invariant statistics $\{T_n\}$ under the scale-change group of transformations for the change-point problem $F(\eta x)$ prechange and $F(\eta\alpha x)$ postchange with unknown $\eta > 0$ and specified $\alpha \neq 1$ is given, for $n+1 \geq k \geq 1$, by*

$$\Lambda_k^n = \alpha^{\beta(n-k+1)}\left[\alpha + (1-\alpha)\prod_{j=k}^{n} T_j\right]^{-n\beta},$$

where $T_1 = 0$ and $T_n = \sum_{j=1}^{n-1} X_j \big/ \sum_{j=1}^{n} X_j$ for $n > 1$.

(ii) *For the invariant Shiryayev–Roberts rule with critical level A,*

$$\lim_{A\to\infty} E_\infty\{N_A\}/A = \Delta(f(x), f(\alpha x)),$$

where $\Delta(f(x), f(\alpha x))$ is the limit corresponding to the fully specified detection rule for $F(x)$ prechange and $F(\alpha x)$ postchange.

PROOF. A maximal invariant for the problem is given by the sequence $Y_j = X_j/X_1$, for $j = 1, 2, \ldots$. Note that $Y_1 = 1$. Write $S_n = \sum_{j=1}^{n} X_j$ and let $n \geq k > 1$. Compute the joint distribution of the first n invariants from

$$P_k\{Y_2 \leq y_2, \ldots, Y_n \leq y_n\}$$
$$= \int_0^\infty \int_0^{x_1 y_2} \cdots \int_0^{x_1 y_n} \prod_{j=1}^{k-1} \frac{x_j^{\beta-1} e^{-x_j}}{\Gamma(\beta)} \prod_{j=k}^{n} \frac{\alpha^\beta x_j^{\beta-1} e^{-\alpha x_j}}{\Gamma(\beta)}\, dx_n \cdots dx_2\, dx_1.$$

Differentiation yields the density and then, for $n \geq k > 1$, the likelihood ratio.

When $k = 1$ or $k = n+1$, the distribution of $X_1, \ldots, X_n$ is identical under P_1 or P_{n+1} to that under P_∞, save perhaps for a different scale. Hence the distribution of any scale-invariant statistic based on the first n observations is the same under P_1 or P_{n+1} as under P_∞, and so $\Lambda_1^n = \Lambda_{n+1}^n = 1$ identically, proving the first assertion of Theorem 1.

The second assertion is proved by verifying the conditions of Theorem 1. Conditions 1A and 1B follow from elementary exponential bounds on tails of gamma-distributed variates. If $\alpha > 1$, Condition 1C is immediate from the boundedness of T_j. The case $\alpha < 1$ is substantially more complicated, requiring study of the conditional distribution of the overshoot of the likelihood ratio, given that the jump over the boundary occurs at a specified time. Monotonicity arguments provide the needed uniform bounds. Full details will be given elsewhere. □

3.2. *Detecting a change in a normal sequence with unknown initial mean.* We next consider a problem studied by Pollak and Siegmund (1991). Suppose that prechange observations are known to be normal with unit variance and mean μ_0. Postchange, the distribution is known to be normal with unit variance and mean $\mu_1 > \mu_0$, where both means are unknown. Here the nuisance parameter is μ_0. This problem is invariant under translation, with maximal invariant $\{T_i = X_i - X_1,\ 2 \le i \le n\}$ at time n. Choose $\delta > 0$ and let the postchange distribution be represented by $\mathscr{N}(\mu_0 + \delta, 1)$. We denote the standard normal density by $\phi(\cdot)$.

THEOREM 3. *Consider the changepoint problem $\mathscr{N}(\mu_0, 1)$ prechange and $\mathscr{N}(\mu_0 + \delta, 1)$ postchange with unknown μ_0 and specified $\delta > 0$.*

(i) *The likelihood ratio of the sequence of statistics maximally invariant under translations is given for $1 \le k \le n+1$ by*

$$\Lambda_k^n = \exp\big((k-1)\big(\delta\big[\overline{X}_n - \overline{X}_{k-1}\big] - \tfrac{1}{2}\delta^2[(n-k+1)/n]\big)\big),$$

where $\overline{X}_n = \sum_{j=1}^n X_j/n$ when $n \ge 1$ and $\overline{X}_0 = 0$.

(ii) *For the invariant Shiryayev–Roberts rule with critical level A, the ARL to false alarm has limit $\lim_{A\to\infty} E_\infty\{N_A\}/A = \Delta(\phi(x), \phi(x-\delta))$, where $\Delta(\phi(x), \phi(x-\delta))$ is the corresponding limit for the fully specified detection rule for $\mathscr{N}(0,1)$ prechange and $\mathscr{N}(\delta, 1)$ postchange.*

PROOF. Under any of the measures P_k the invariants $T_2, \dots, T_n$ are multivariate normal with identical covariance matrix $I + J$, where J is the matrix having all entries 1, yielding the first assertion.

To prove the second assertion, we use Theorem 1. We begin with Condition 1A. Let $0 < \varepsilon_1, \varepsilon_2 < 1$ be given. Write $\varepsilon_3 = (\varepsilon_1(1-\varepsilon_1) \wedge \varepsilon_2(1-\varepsilon_2))/2$ and set $a_1 = \delta^2\varepsilon_3/4$. Suppose $k - 1 \in [n\varepsilon_1, n(1-\varepsilon_2)]$. For $n \ge 1/\varepsilon_3$,

$$P_\infty\{\Lambda_k^n < e^{-a_1 n}\} = P_\infty\left\{(k-1)\left(\delta\big[\overline{X}_n - \overline{X}_{k-1}\big] - \frac{\delta^2}{2}\left[\frac{n-k+1}{n}\right]\right) < -\frac{\delta^2\varepsilon_3 n}{4}\right\},$$

from which Condition 1A follows by standard bounds on the tails of the normal distribution.

We now verify Condition 1B. Because the likelihood ratio Λ_k^n is invariant, we may take $\mu_0 = 0$ with no loss of generality. Hence $f_1(x)/f_0(x) = \exp(\delta x -$

$\delta^2/2)$. Write $\theta = (n-k+2)/(n+1)$, so that

$$\left|\log\left(\frac{\Lambda_k^{n+1}}{\Lambda_k^n}\Big/\frac{f_1(X_{n+1})}{f_0(X_{n+1})}\right)\right| < \delta\theta|X_{n+1}| + \delta^2\left(\theta + \frac{1}{n+1}\right) + \delta|\overline{X}_n|,$$

which is easily used to establish Condition 1B.

Lastly, we consider Condition 1C. Let $\mathscr{F}_n$ be the σ-field generated by the first $n-1$ invariant statistics $X_2 - X_1, \dots, X_n - X_1$. Follow exactly the lines of Pollak and Siegmund's (1991) Lemma 2, replacing ν_0 by an arbitrary stopping time M. We thus obtain a constant k_1 such that for all $A > 1$ and all stopping times M,

$$E_\infty\left\{\left[\sum_{k=M+1}^{N}\Lambda_k^N\right]^2 \,\Big|\, \mathscr{F}_M\right\} \le (1+A)^2 E_\infty\left\{k_1 + \exp\left(2\delta^2 + 2\delta\max_{i\ge M}|\overline{X}_i|\right) \,\Big|\, \mathscr{F}_M\right\},$$

where $N = N_A^M$. Observe that $\overline{X}_n$ is independent of the invariant σ-field $\mathscr{F}_n$ for each integer n. Hence, using the strong Markov property implies that $E_\infty\{[\sum_{k=M+1}^N \Lambda_k^N]^2 \mid \mathscr{F}_M\} \le k_2 A^2$, for some positive constant k_2 and all $A > 1$. Now apply the alternate condition given after Theorem 1. □

3.3. *Detecting a shift to a stochastically larger distribution.* Our final example is the nonparametric problem which motivated the formulation of our main theorem. The ARL to false alarm of parametric cusum procedures is known to be quite sensitive to small perturbations of the prechange distribution. This nonrobustness is of substantial practical concern. See, for example, van Dobben de Bruyn [(1968), Section 2.4]. Use of a procedure based on ranks—which is invariant to strictly increasing transformations—is a natural response to this problem. Specifically, the procedure described below in Theorem 4 permits us to assert known ARL to false alarm regardless of the prechange distribution. Such robustness of validity can be purchased at small price. For the $\mathscr{N}(\mu, 1)$ prechange versus $\mathscr{N}(\mu+1, 1)$ postchange detection problem, calculations similar to those of Gordon and Pollak (1994) show that choosing $p = 0.8413$, $\alpha = 0.53$ and $\beta = 1.7$ yields a procedure with 97% asymptotic relative efficiency as $A \to \infty$.

Although the procedure is computation intensive, it is still feasibly implemented with current technology. A complete analysis of the operating characteristics of the procedure, including a proof of Theorem 4 by verifying the hypotheses of Theorem 1 is the subject of Gordon and Pollak (1995).

THEOREM 4. *Consider the changepoint problem where* $f_0(x) = e^{-|x|}/2$ *prechange, and* $f_1(x) = p\alpha e^{-\alpha x} I_{\{x>0\}} + q\beta e^{\beta x} I_{\{x\le 0\}}$ *postchange. If* $p > 1/2$, $\alpha \le 1 \le \beta$ *and* $p\alpha \ge q\beta$ *then*:

(i) *The ranks are invariant under the group of strictly increasing transformations of the data. Given the permutation* $\rho(\cdot, n)$ *determined by the ranks with inverse permutation* $\tau(\cdot, n)$, *the corresponding likelihood ratio function* *for* $1 \le k \le n+1$ *is*

$$\Lambda_k^n(\rho(\cdot, n)) = \frac{P_{\nu=k}\{X_{\tau(1,n)} < \cdots < X_{\tau(n,n)}\}}{P_{\nu=\infty}\{X_{\tau(1,n)} < \cdots < X_{\tau(n,n)}\}} = \sum_{m=0}^{n} \lambda_{k,m}^n(\rho(\cdot, n)),$$

where $U_k(i,n) = \sum_{j=k}^n I_{\{\rho(j,n)>i\}}$, $V_k(i,n) = (n+1-k) - U_k(i,n)$ *and*

$$\lambda_{k,m}^n(\rho(\cdot,n)) = \binom{n}{m}\left(\frac{1}{2}\right)^n \left(\frac{p\alpha}{q\beta}\right)^{U_k(m,n)} (2q\beta)^{n+1-k}$$

$$\times \prod_{i=1}^{m}\left(1 + \frac{V_k(i,n)}{i}(\beta - 1)\right)^{-1} \prod_{i=m+1}^{n}\left(1 + \frac{U_k(i-1,n)}{n+1-i}(\alpha-1)\right)^{-1}.$$

(ii) *Given the Shiryayev–Roberts stopping rule* $N_A = \inf\{n \mid \sum_{k=1}^n \Lambda_k^n \ge A\}$, *its expected time to false alarm has limit* $\lim_{A\to\infty} E_\infty\{N_A/A\} = \Delta(f_0, f_1)$ *where* $\Delta(f_0, f_1)$ *is the same limiting value as obtains for the parametric Shiryayev–Roberts procedure.*

4. Proof of Theorem 1. We follow the structure of Pollak (1987). The proof is divided into a series of lemmas, with intervening discussion. To facilitate reference, we begin with a listing of notation which will be reintroduced when needed.

4.1. *Notation.* We write $\Delta = \Delta(f_0, f_1)$, the constant corresponding to the detection problem with nuisance parameter known. Let $Z_i = \log(f(X_i|\eta_0, \alpha(\eta_0))/f_0(X_i|\eta_0))$ denote the log-likelihood ratio for the fully parameterized problem. We make the critical assumption that the distribution of the Z_i is continuous.

Let $c > 1$ be a large integer whose value is fixed in Section 4.3. From Jensen's inequality, $E_\infty\{Z_i\} < 0$, so we let n_c be the least positive integer with $P_\infty\{\exp(\sum_{i=1}^{n_c} Z_i) > c^{-8}\} < c^{-4}$. Note that n_c is nondecreasing in c.

Inductively define times L_j adapted to the invariant σ-fields $\mathscr{F}_n$. Let $L_0 = 0$. For $j > 0$,

$$L_j = \min\left\{n \,\middle|\, n > L_{j-1} \text{ and } \sum_{k=L_{j-1}+1}^{n} \Lambda_k^n \ge A/c\right\}.$$

Let $W = \max\{j \mid L_j \le N_A\}$. Note that W is not a stopping time.

The contributions to R_n indexed by times after L_{j-1} are, for $j \ge 1$,

$$Q(j,n) = \begin{cases} \displaystyle\sum_{k=L_{j-1}+1}^{n} \Lambda_k^n, & \text{if } n > L_{j-1}, \\ 0, & \text{otherwise.} \end{cases}$$

Next define the fully parameterized continuation of $Q(j, L_j)$,

$$V(j,n) = \begin{cases} Q(j,n), & \text{if } n \le L_j, \\ Q(j,L_j)\exp\left(\displaystyle\sum_{i=L_j+1}^{n} Z_i\right), & \text{if } n > L_j, \end{cases}$$

and the invariant continuation of $Q(j, L_j)$,

$$v(j,n) = \begin{cases} Q(j,n), & \text{if } n \le L_j, \\ \displaystyle\sum_{k=L_{j-1}+1}^{L_j} \Lambda_k^n, & \text{if } n \ge L_j. \end{cases}$$

Both $V(j,n)$ and $v(j,n)$ are sums of $L_j - L_{j-1}$ terms when $n \ge L_j$.

Corresponding to the fully specified continuation $V(j,n)$, define stopping times

$$H_j = \begin{cases} \min\{n \mid n \ge L_j \text{ and } V(j,n) \ge A\}, \\ \infty \text{ if the set of such } n \text{ is empty,} \end{cases}$$

$$M_j = H_j \wedge (L_j + n_c) \wedge L_{j+1}$$

with analogous quantities for the invariant continuation,

$$h_j = \begin{cases} \min\{n \mid n \ge L_j \text{ and } v(j,n) \ge A\}, \\ \infty \text{ if the set of such } n \text{ is empty,} \end{cases}$$

$$m_j = h_j \wedge (L_j + n_c) \wedge L_{j+1}.$$

Finally, define random indices

$$J = \min\{j \mid v(j, m_j) \ge A\},$$
$$J^* = J \wedge (c^3 + 1),$$

and, analogously, given a stopping time Y, let

$$J_Y = \min\Big\{j \mid \sup_{n \ge Y} v(j,n) \ge A\Big\},$$
$$J_Y^* = J_Y \wedge (c^3 + 1).$$

As is usual, given a stopping time M, we write $\mathscr{F}_M$ for the σ-field associated with the stopping time. We frequently indicate conditioning on $\mathscr{F}_M$ by explicitly conditioning on $M; T_1, \dots, T_M$.

Write $\hat{o}_c(A)$ for functions $\psi(A,c)$ such that $\limsup_{A\to\infty} \big| \psi(A,c)/A \big|$ can be made arbitrarily small by choosing and fixing c large. Similarly, we write $\hat{o}_c(1)$ for functions $\psi(A,c)$ such that $\limsup_{A\to\infty} \big| \psi(A,c) \big|$ can be made small by fixing c large.

Our plan is to show that R_{N_A} is well approximated by $v(J^*, m(J^*))$. The idea is that when R_n becomes large—say of the order A/c—then Condition 1A tells us that the summands contributing by far the greatest amount are those Λ_k^n with k close to n. Hence for n's not much larger than some L_j, we have $R_n \approx v(j,n)$ from which $R_{N_A} \approx v(J, m(J))$ will follow. We will show that with high probability $\{v(j,n)\}_{n=L_j}^{\infty}$ either crosses the level A quickly or never crosses A. Therefore, because of Condition 1B, and because the summands which contribute to $v(j,n)$ have k's close to n, the overshoot of $v(j,n)$ over A—if it crosses at all—should resemble closely the overshoot in the fully specified problem. Hence renewal theory will yield the limit $E_\infty\{R_{N_A}\}/A \to \Delta$, the same limit as in the fully specified problem.

Lemmas 1 to 3 give bounds related to the increments in stopping times referred to in Condition 1C. Lemma 4 states that the number of L_j's observed up to time cA is bounded by c^3 with high probability. Lemma 5 states that the time between successive L_j's tends to be large.

Lemma 6 formalizes the idea that $v(j,n) \approx V(j,n)$ for n only slightly larger than L_j, while Lemma 7 tells us that $v(j,n)$ cannot be large for n much larger than L_j. These imply Lemma 8 which states that N_A is close to one of the first c^3 stopping times L_j. Lemmas 9 to 14 allow us to approximate the expectations of $v(j, m_j)$ conditional on $v(j, m_j) \geq A$, in order to obtain $E\{v(J^*, m_{J^*})\}/A \approx \Delta$. The rest of the proof shows that $E_\infty\{R_{N_A} - v(J^*, m_{J^*})\}/A$ is negligible, which finally yields Theorem 1.

4.2. *Preliminary lemmas.*

LEMMA 1. *For stopping times M and $N = N_A^M$ as in Condition* 1C,

$$E_\infty\{N_A^M - M \mid M; T_1, \dots, T_M\} \leq \kappa(1)A.$$

In particular, $E_\infty\{N_A\} \leq \kappa(1)A$ *and* $E_\infty\{L_{j+1} - L_j\} \leq \kappa(1)A/c$, *for* $j \geq 0$.

PROOF. The centered sum of likelihood ratios $\sum_{k=M+1}^{M+n} \Lambda_k^{M+n} - n$ is a P_∞-martingale adapted to $\{M; T_1, \dots, T_{M+n}\}$ for $n \geq 0$. Truncate the stopping time N_A^M at some large finite time $M + m$ to obtain from Condition 1C that $E_\infty\{(N - M) \wedge m \mid M; T_1, \dots, T_M\} \leq \kappa(1)A$. Now let $m \to \infty$. □

LEMMA 2. *If M and $N = N_A^M$ are as in Condition* 1C, *then conditional on the entire past through M, the distribution of $N - M$ is stochastically larger than a uniform*$[0, A]$ *random variate. Specifically, for all $t \in [0, 1]$,*

$$P_\infty\{N - M \leq tA \mid M; T_1, \dots, T_M\} \leq t.$$

PROOF. Apply the optional sampling theorem to $\sum_{k=M+1}^{M+n} \Lambda_k^{M+n} - n$ yielding

$$tA \geq E_\infty\left\{ \sum_{k=M+1}^{M+((N-M)\wedge\lfloor At\rfloor)} \Lambda_k^{M+((N-M)\wedge\lfloor At\rfloor)} \,\middle|\, \mathscr{F}_M \right\} \geq A\, P_\infty\{N - M \leq At \mid \mathscr{F}_M\}. \quad \square$$

LEMMA 3. *If M and $N = N_A^M$ are as in Condition* 1C, *then for all t positive*,

$$P_\infty\{N - M \geq At \mid M; T_1, \dots, T_M\} \leq \kappa(1)/t.$$

In particular, $P_\infty\{N - M \geq At\} \leq \kappa(1)/t$.

PROOF. Lemma 1 and Markov's inequality. □

Recall that we have partitioned the observation times into exceedance epochs L_j. Hence we have the decomposition

$$R_n = \sum_{j=1}^{n} \sum_{k=L_{j-1}+1}^{L_j \wedge n} \Lambda_k^n,$$

where we follow the usual convention that summation over a null set of indices yields a sum with value 0.

LEMMA 4. *If* $c \ge 5$, *then* $P_\infty\{L_{c^3} < cA\} < c^{-3}$.

PROOF. From Lemma 2, the sum $\sum_{i=1}^{c^3}(L_i - L_{i-1})$ is stochastically greater than the sum $\sum_{i=1}^{c^3} U_i$, where the U_i are i.i.d. uniform$(0, A/c)$. The former sum is exactly L_{c^3}. The bound follows from Chebychev's inequality applied to $\sum_{i=1}^{c^3} U_i$. □

LEMMA 5. $P_\infty\{L_j - L_{j-1} > Ac^{-5}$ *for all* $1 \le j \le c^3\} \ge 1 - c^{-1}$.

PROOF. From Lemma 2, $P_\infty\{L_j - L_{j-1} \le \varepsilon A/c | \mathscr{F}_{L_{j-1}}\} \le \varepsilon$. Now set $\varepsilon = c^{-4}$ and sum over j. □

In the following, we write $\overline{B}$ for the complement of an event B.

LEMMA 6. *Given* $0 < \omega_1, \omega_2 < 1$ *and* $c > 4$, *there exist positive functions* $A^{(6)} = A^{(6)}(c, \omega_1, \omega_2)$ *and* $b^{(6)} = b^{(6)}(c, \omega_1, \omega_2)$ *such that*

$$P_\infty\left\{\sup_{1\le i\le 2n_c}\left|\frac{v(j, L_j + i)}{V(j, L_j + i)} - 1\right| > \omega_1 \text{ and } L_1 > \frac{A}{c^4}\,\middle|\,\mathscr{F}_{L_{j-1}}\right\} < \omega_2 + D_{j-1}$$

whenever $j > 1$ *and* $A > A^{(6)}$, *where* $D_{j-1} \in \mathscr{F}_{L_{j-1}}$ *is a Bernoulli variate with* $E_\infty\{D_{j-1}\} < \exp(-b^{(6)}A)$.

PROOF. Whenever unambiguous, we suppress the explicit dependence of various quantities upon the fixed values c, ω_1, and ω_2. Choose and fix $\alpha(c, \omega_1, \omega_2)$ large so that

$$P_\infty\left\{\sup_{1\le i\le 2n_c}\exp\left(-\sum_{l=1}^{i} Z_l\right) > \frac{\alpha\omega_1}{2}\right\} < \frac{\omega_2}{12}. \tag{1}$$

Let $B_0 = \bigcup_{i=1}^{2n_c}\{X_{L_j+i} \notin B_\varepsilon\}$, where B_ε is as in Condition 1B. Clearly, $P_\infty\{B_0\} \le 2n_c\varepsilon$. Choose ε small so that $P_\infty(B_0) \le \omega_2/2$.

Let M be a stopping time and let $1 \le m_0 < \infty$. By Condition 1B,

$$\begin{aligned} P_\infty\Bigg\{\max_{(1-\theta)M\le k\le M+1}\left|1 - \frac{\Lambda_k^{M+1}}{\Lambda_k^M} \Big/ \frac{f_1(X_{M+1})}{f_0(X_{M+1})}\right| > \varepsilon \\ \text{and } X_{M+1} \in B_\varepsilon \text{ and } M \ge m_0\Bigg\} \\ \le \frac{b_1}{1 - \exp(-b_2)}\exp(-b_2 m_0). \end{aligned} \tag{2}$$

The times L_j are monotone in j, so $\mathscr{F}_{L_{j-1}} \subset \mathscr{F}_{L_j}$. Apply (2) $2n_c$ times to obtain $n_0^*(c, \omega_1, \omega_2)$, $\theta^*(c, \omega_1, \omega_2)$ positive and $b^*(c, \omega_1, \omega_2)$ such that for

$$B_1 = \left\{ \sup_{1 \le i \le 2n_c} \sup_{(1-\theta^*)L_j \le k \le L_j} \left| \frac{\Lambda_k^{L_j+i}}{\Lambda_k^{L_j} \exp(\sum_{l=L_j+1}^{L_j+i} Z_l)} - 1 \right| > \frac{\omega_1}{2} \right\}$$

we have $P_\infty\{\overline{B}_0 \text{ and } B_1 \text{ and } L_1 > n\} < \exp(-b^* n)$, whenever $n > n_0^*$. Now consider the event $B_2 = \{\sup_{0 \le i \le 2n_c} \sum_{\alpha^{-1}L_j \le k \le (1-\theta^*)L_j} \Lambda_k^{L_j+i} > c^{-1}\}$, the event $B_3 = \{L_{j-1} < \alpha^{-1} L_j\}$ and $B_4 = \{L_1 > A/c^4\}$. Use Condition 1A to choose a_1^*, a_2^* and a_3^* such that $P_\infty\{\sup_{0 \le i \le 2n_c} \sum_{\alpha^{-1}n \le k \le (1-\theta^*)n} \Lambda_k^{n+i} > n \exp(-a_1^* n)\} \le a_2^* \exp(-a_3^* n)$. Hence there exist $n_0^{**}(c, \omega_1, \omega_2) > n_0^*$ and $b^{**}(c, \omega_1, \omega_2) > 0$ such that

$$(3) \qquad P_\infty\{\overline{B}_0 \text{ and } B_1 \text{ and } L_1 > n\} + P_\infty\{B_2 \text{ and } L_1 > n\} < \omega_2 \exp(-b^{**} n)/12,$$

for $n \ge n_0^{**}$. Let $A^{(6)}(c, \omega_1, \omega_2) = (n_0^{**}\alpha) \vee (2/\omega_1)$ and assume $A > A^{(6)}$. By definition, $\{L_1 \ge n_0^{**}\} \supset B_4$. Use Condition 1C and Lemma 2 to show that

$$(4) \qquad P_\infty\{B_3 \,|\, \mathscr{F}_{L_{j-1}}\} = P_\infty\{(\alpha - 1)L_{j-1} < L_j - L_{j-1} \,|\, \mathscr{F}_{L_{j-1}}\} \le \frac{A/c}{3c^4 L_1} \omega_2.$$

Now let $\Psi_{j-1} = P_\infty\{(\overline{B}_0 B_1 \cup B_2) B_4 \,|\, \mathscr{F}_{L_{j-1}}\}$. From (3), $E_\infty\{\Psi_{j-1}\} < (\omega_2/12)\exp(-b^{**}A/c^4)$. Let D_{j-1} be the indicator of the event $\{\Psi_{j-1} > \omega_2/12\}$. Apply the Markov inequality to choose $b^{(6)}$ for which

$$(5) \qquad E_\infty\{D_{j-1}\} < \exp(-b^{(6)} A),$$

increasing A_0, if necessary. Recall $A > \alpha$ and $Q(j, L_j) \ge A/c$. Use (1), (3), (4) and (5) to obtain on B_4 that

$$P_\infty\left\{ B_4 \text{ and } \overline{B}_0 \text{ and } \sup_{1 \le i \le 2n_c} \left| \frac{v(j, L_j + i)}{V(j, L_j + i)} - 1 \right| > \omega_1 \,\middle|\, \mathscr{F}_{L_{j-1}} \right\} < \frac{\omega_2}{2} + D_{j-1}.$$

Combine $P_\infty\{B_0\} \le \omega_2/2$ with the preceding to obtain the bound $\omega_2 + D_{j-1}$, proving the lemma. □

Recall that Z_i are the fully specified increments in log-likelihood ratio. Prechange, they are i.i.d. with negative drift. We now use the defining property of n_c, that $P_\infty\{\exp(\sum_{i=1}^{n_c} Z_i) > c^{-8}\} < c^{-4}$.

LEMMA 7. *There exists a function $A^{(7)}(c)$ and constants $c^{(7)}$ and $k^{(7)}$ such that both*

$$P_\infty\left\{ \max_{2 \le j \le c^3} \max_{n \ge n_c} v(j, L_j + n) > A/c \right\} < \frac{k^{(7)}}{c}$$

and

$$P_\infty\left\{ \max_{2 \le j \le c^3} \max_{n \ge 2n_c} v(j, L_j + n) > A/c^9 \right\} < \frac{k^{(7)}}{c},$$

whenever $c > c^{(7)}$ and $A > A^{(7)}(c)$.

PROOF. Let $B_0 = \{L_1 \le A/c^4\}$. Define

$$B_1(j) = \left\{ \frac{v(j, L_j + n_c)}{V(j, L_j + n_c)} > 1 + \frac{1}{c} \right\} \quad \text{and} \quad B_2(j) = \{V(j, L_j + n_c) > c^{-5}A\}$$

for $2 \le j \le c^3$. Use Lemma 6 to choose $A^{(7)}(c) \ge A^{(6)}(c, c^{-1}, c^{-4}/2)$ sufficiently large that $A > A^{(7)}$ implies

$$P_\infty\{\overline{B}_0 B_1(j)\} < \frac{1}{c^4}. \tag{6}$$

Also note that

$$\begin{aligned} P_\infty\{B_2(j)\} &\le P_\infty\{Q(j, L_j) > c^3 A\} \\ &\quad + P_\infty\left\{Q(j, L_j) \le c^3 A \text{ and } \exp\left(\sum_{l=1}^{n_c} Z_{L_j+l}\right) > c^{-8}\right\} \\ &\le \frac{\kappa(1)}{c^4} + \frac{1}{c^4}. \end{aligned} \tag{7}$$

We may use the martingale maximum inequality and Condition 1C to bound $P_\infty\{\max_{n \ge n_c} v(j, L_j + n) > A/c;\ \overline{B_0 B_1(j) B_2(j)}\}$ because $v(j, L_j + n)$ is a martingale for $n \ge 0$. From Lemma 2, $P_\infty\{B_0\} < c^{-3}$. Use (6) and (7) to complete the proof of the first assertion.

Prove the second similarly by replacing n_c with $2n_c$ in the definitions of $B_1(j)$ and $B_2(j)$ and bounding $P_\infty\{V(j, L_j + 2n_c) > c^{-13}A\}$. □

Recall that $W = \max\{j | L_j \le N_A\}$ and that W is not a stopping time.

LEMMA 8. *There exists a function $A^{(8)}(c)$ and constants $c^{(8)}$ and $k^{(8)}$ such that*

$$P_\infty\{N_A - L_W > n_c\} \le \frac{k^{(8)}}{c}$$

and

$$P_\infty\{W \ge c^3\} \le \frac{k^{(8)}}{c},$$

whenever $c > c^{(8)}$ and $A > A^{(8)}(c)$.

PROOF. Take $c > 5$ and choose

$$A^{(8)}(c) = A^{(6)}(c, c^{-4}, c^{-4}) \vee [4 \log(c)/b^{(6)}(c, c^{-4}, c^{-4})].$$

Let $A > A^{(8)}(c)$. Observe

$$A \le R_{N_A} = v(1, N_A) + \sum_{j=2}^{W-1} v(j, N_A) + \sum_{k=L_{W-1}+1}^{L_W} \Lambda_k^{N_A} + \sum_{k=L_W+1}^{N_A} \Lambda_k^{N_A}.$$

Decompose the event of interest as

$$P_\infty\{N_A - L_W > n_c\} \le P_\infty\{W \ge c^3\} \tag{8}$$

$$+ P_\infty\{v(1, N_A) > A/4\} \tag{9}$$

$$+ P_\infty\left\{\sum_{j=2}^{W-1} v(j, N_A) > A/4 \text{ and } W < c^3\right\} \tag{10}$$

$$+ P_\infty\left\{\max_{2\le j\le c^3}\max_{n\ge n_c} v(j, L_j + n) > A/4\right\} \tag{11}$$

$$+ P_\infty\left\{\sum_{k=L_W+1}^{N_A} \Lambda_k^{N_A} > A/4\right\}. \tag{12}$$

Use Lemmas 3 and 4 to show

$$P_\infty\{W \ge c^3\} \le P_\infty\{L_{c^3} < cA\} + P_\infty\{N_A \ge cA\} \le c^{-3} + \kappa(1)/c,$$

bounding (8) and proving the second assertion of the lemma.

Bound (9) by $4\kappa(1)/c$ using the martingale maximum inequality and Condition 1C. Use Lemmas 2, 5, 6, the choice of $A > A^{(8)}(c)$ and the martingale maximum inequality to bound (10) by $9/c$. Next, bound (11) using Lemma 7. Finally, (12) has probability 0, by the definition of W as a maximum. □

Recall that h_j is the first time after the stopping time L_j that the invariant continuation of $\sum_{k=L_{j-1}+1}^{L_j} \Lambda_k^{L_j}$ is at least as large as A. The next lemma tells us that if we do go over the barrier A at the jth excursion above A/c, then with high probability the exceedance occurs very soon after L_j.

LEMMA 9. *Let $j > 1$. There exist positive functions $A^{(9)}(c)$ and $b^{(9)}(c)$ such that $A > A^{(9)}(c)$ and $c > 4$ implies*

$$P_\infty\{L_j + n_c < h_j < \infty \text{ and } L_1 > A/c^4 \mid \mathscr{F}_{L_{j-1}}\} < 3c^{-4} + D_{j-1},$$

where $D_{j-1} \in \mathscr{F}_{L_{j-1}}$ is a Bernoulli variate with $E_\infty\{D_{j-1}\} < \exp(-b^{(9)}A)$.

PROOF. Note that $L_j < h_j$ implies $Q(j, L_j) < A$. Hence, on $\{L_1 > A/c^4\}$ we may bound the conditional probability of interest by

$$P_\infty\left\{\frac{v(j, L_j + n_c)}{V(j, L_j + n_c)} > 2 \,\middle|\, \mathscr{F}_{L_{j-1}}\right\} \tag{13}$$

$$+ P_\infty\left\{\frac{V(j, L_j + n_c)}{Q(j, L_j)} > c^{-7} \,\middle|\, \mathscr{F}_{L_{j-1}}\right\} \tag{14}$$

$$+ P_\infty\left\{L_j + n_c < h_j < \infty \text{ and } \frac{v(j, L_j + n_c)}{A} \le 2c^{-7} \,\middle|\, \mathscr{F}_{L_{j-1}}\right\}. \tag{15}$$

For A large enough, we bound (13) using Lemma 6. We bound (14) using the strong Markov property for i.i.d. variates with the definitions of n_c and $V(j, L_j + n_c)$. Finally, use the martingale maximum inequality to bound (15) by $2c^{-7}$. □

LEMMA 10. *Let $j > 1$. There exist positive functions $c^{(10)}(\xi)$, $b^{(10)}(c, \xi)$ and $A^{(10)}(c, \xi)$ and a positive constant $k^{(10)}$ such that, given $0 < \xi < 1$, on the event $\{L_1 > A/c^4\}$,*

$$\left| \frac{E_\infty\{Q(j, L_j) \mid \mathscr{F}_{L_{j-1}}\}}{A\Delta} - P_\infty\left\{h_j < \infty \mid \mathscr{F}_{L_{j-1}}\right\} \right| < k^{(10)}\left(\frac{\xi}{c} + D_{j-1}\right),$$

where $D_{j-1} \in \mathscr{F}_{L_{j-1}}$ is a Bernoulli variate with $E_\infty\{D_{j-1}\} < \exp(-b^{(10)}A)$ whenever $c > c^{(10)}(\xi)$ and $A > A^{(10)}(c, \xi)$.

PROOF. Recall that Z_i has positive mean under P_{L_j+1} and that its distribution is continuous. Because $V(j, L_j + n)/Q(j, L_j)$ is a likelihood ratio,

$$\begin{aligned} &P_\infty\left\{H_j < \infty \mid \mathscr{F}_{L_j}\right\} \\ &\quad = P_\infty\left\{\sup_{n\ge 0} \sum_{i=L_j+1}^{L_j+n} Z_i \ge \log(A) - \log(Q(j, L_j)) \,\Big|\, \mathscr{F}_{L_j}\right\} \\ &\quad = E_{L_j+1}\left\{\exp\left(-\left[\sum_{i=L_j+1}^{H_j} Z_i\right] + \log(A/Q(j, L_j))\right) \,\Big|\, \mathscr{F}_{L_j}\right\} \frac{Q(j, L_j)}{A}. \end{aligned} \tag{16}$$

Given fixed $\xi > 0$, we may use renewal theory as in Siegmund (1985), Chapter VIII to approximate the expectation in (16) by $1/\Delta = 1/\Delta(f_0, f_1)$ with error less than ξ/Δ whenever $\log(A) - \log(Q(j, L_j))$ exceeds some threshold $\tau(\xi)$. Hence, Condition 1C and Markov's inequality imply

$$\begin{aligned} &P_\infty\left\{\log(A) - \log(Q(j, L_j)) \le \tau(\xi) \mid \mathscr{F}_{L_{j-1}}\right\} \\ &\quad \le \frac{1}{c}\exp(\tau(\xi)) E_\infty\left\{\frac{Q(j, L_j)}{A/c} I_{\{(A/c)^{-1}Q(j,L_j)\ge c\exp(-\tau(\xi))\}} \,\Big|\, \mathscr{F}_{L_{j-1}}\right\} \\ &\quad \le \frac{1}{c}\exp(\tau(\xi))\kappa(c\exp(-\tau(\xi))), \end{aligned} \tag{17}$$

showing also that $E_\infty\{A^{-1}Q(j, L_j)I_{\{A^{-1}Q(j,L_j)\ge\exp(-\tau(\xi))\}} \mid \mathscr{F}_{L_{j-1}}\}$ is bounded by (17). Because $\kappa(t) \to 0$ as $t \to \infty$, we may choose $c^{(10)}(\xi)$ such that $c > c^{(10)}(\xi)$ implies both $\exp(\tau(\xi))\kappa(c\exp(-\tau(\xi))) < \xi$ and $1/c < \xi$.

Let $c > c^{(10)}(\xi)$. From (16) and renewal theory,

$$\begin{aligned} P_\infty\left\{H_j < \infty \mid \mathscr{F}_{L_j}\right\} &\ge P_\infty\left\{H_j < \infty \text{ and } \frac{Q(j, L_j)}{A} < \exp(-\tau(\xi)) \,\Big|\, \mathscr{F}_{L_j}\right\} \\ &\ge (1 - \xi)\frac{Q(j, L_j)}{A\Delta}\left(1 - I_{\{Q(j,L_j)/A\ge\exp(-\tau(\xi))\}}\right). \end{aligned}$$

Now take expectations with respect to the coarser σ-field $\mathscr{F}_{L_{j-1}}$ and apply (17), Condition 1C and the definition of $c^{(10)}(\xi)$ to obtain

$$(18)\qquad P_\infty\{H_j<\infty\,|\,\mathscr{F}_{L_{j-1}}\}\ge E_\infty\left\{\frac{Q(j,L_j)}{A\Delta}\,\middle|\,\mathscr{F}_{L_{j-1}}\right\}-\frac{\xi(\kappa(1)+1)}{c\Delta}.$$

Similarly, again using (17),

$$(19)\qquad P_\infty\{H_j<\infty\,|\,\mathscr{F}_{L_{j-1}}\}\le E_\infty\left\{\frac{Q(j,L_j)}{A\Delta}\,\middle|\,\mathscr{F}_{L_{j-1}}\right\}+\frac{\xi}{c}\left(\frac{\kappa(1)}{\Delta}+1\right).$$

It remains to approximate h_j with H_j. Use the continuity of the distribution of Z_i and the finiteness of n_c to choose $\delta=\delta(c,\xi)\in(0,1/2)$ such that $\max_{x\ge 0}\sum_{j=1}^{n_c}P_\infty\{\sum_{i=1}^{j}Z_i\in(x-\delta,x+\delta)\}<\xi/c$. Define events

$$B_0=\{L_1\le A/c^4\}$$

$$B_1=\left\{\min_{1\le i\le n_c}\left|\left[\log(Q(j,L_j))+\sum_{l=1}^{i}Z_{L_j+l}\right]-\log(A)\right|<\delta\right\}$$

$$B_2=\left\{\max_{1\le i\le n_c}\left|\log(v(j,L_j+i)/V(j,L_j+i))\right|>\delta\right\}$$

$$B_3=\{H_j\ne h_j \text{ and } H_j\wedge h_j\le L_j+n_c\}.$$

Conclude from the choice of δ that $P_\infty\{B_1\,|\,\mathscr{F}_{L_{j-1}}\}<\xi/c$. Use Lemma 6 to choose $A^*(c,\delta,\xi)$ and $b^*(c,\delta,\xi)>0$ such that $A>A^*$ implies $P_\infty\{B_2\overline{B}_0\,|\,\mathscr{F}_{L_{j-1}}\}<\xi/c+D^*_{j-1}$, where $D^*_{j-1}\in\mathscr{F}_{L_{j-1}}$ is Bernoulli having $E_\infty\{D^*_{j-1}\}<e^{-b^*A}$.

Use Lemma 9 and the definition of n_c to obtain A^{**} and $b^{**}>0$ for which $A>A^{**}$ implies $|P_\infty\{H_j<\infty\,|\,\mathscr{F}_{L_{j-1}}\}-P_\infty\{h_j<\infty\,|\,\mathscr{F}_{L_{j-1}}\}|$ is bounded above by

$$5c^{-4}+D^{**}_{j-1}+P_\infty\{B_1\,|\,\mathscr{F}_{L_{j-1}}\}+P_\infty\{B_2\,|\,\mathscr{F}_{L_{j-1}}\}+P_\infty\{B_3\overline{B}_2\overline{B}_1\,|\,\mathscr{F}_{L_{j-1}}\}$$

on $\overline{B}_0$ with $D^{**}_{j-1}\in\mathscr{F}_{L_{j-1}}$ Bernoulli having $E_\infty\{D^{**}_{j-1}\}<e^{-b^{**}A}$. On $\overline{B}_1\overline{B}_2$, we have $h_j=H_j$ if $L_j<h_j\wedge H_j\le L_j+n_c$. Therefore $|P_\infty\{H_j<\infty\,|\,\mathscr{F}_{L_{j-1}}\}-P_\infty\{h_j<\infty\,|\,\mathscr{F}_{L_{j-1}}\}|$ is bounded above by $5\xi/c^3+D^{**}_{j-1}+2\xi/c+D^*_{j-1}$ on the event $\overline{B}_0$. Finally, combine this last bound with (18) and (19). □

Recall that the stopping time m_j is defined in terms of the invariant continuation $v(j,\cdot)$ of $Q(j,L_j)$. Specifically, $m_j=h_j\wedge(L_j+n_c)\wedge L_{j+1}$.

LEMMA 11. *There exist a constant $k^{(11)}$ and positive functions $c^{(11)}(\xi)$, $b^{(11)}(c,\xi)$ and $A^{(11)}(c,\xi)$ such that, given $0<\xi<1$, on the event $\{L_1>A/c^4\}$,*

$$\left|\frac{E_\infty\{Q(j,L_j)\,|\,\mathscr{F}_{L_{j-1}}\}}{A\Delta}-P_\infty\{v(j,m_j)\ge A\,|\,\mathscr{F}_{L_{j-1}}\}\right|<k^{(11)}\left(\frac{\xi}{c}+D_{j-1}\right),$$

where $D_{j-1} \in \mathscr{F}_{L_{j-1}}$ *is a Bernoulli variate with* $E_\infty\{D_{j-1}\} < \exp(-b^{(11)}A)$ *whenever* $j > 1$, $c > c^{(11)}(\xi)$ *and* $A > A^{(11)}(c, \xi)$.

PROOF. From Lemma 2, $P_\infty\{L_{j+1} - L_j \le n_c \mid \mathscr{F}_{L_j}\} \le n_c(A/c)^{-1}$. Now condition on the coarser σ-field $\mathscr{F}_{L_{j-1}}$, and then apply Lemmas 9 and 10. □

LEMMA 12. *There exist* $A^{(12)}(c)$ *and* $b^{(12)}(c) > 0$ *for which, on* $\{L_1 > A/c^4\}$,

$$E_\infty\{v(j, m_j);\, v(j, m_j) < A \mid \mathscr{F}_{L_{j-1}}\} < cn_c + 3Ac^{-4} + AD_{j-1},$$

where $D_{j-1} \in \mathscr{F}_{L_{j-1}}$ *is a Bernoulli variate with* $E_\infty\{D_{j-1}\} < \exp(-b^{(12)}A)$ *whenever* $j > 1$, $c > 4$ *and* $A > A^{(12)}(c)$.

PROOF. Consider the events $B_1 = \{Q(j, L_j) < A\}$, $B_2 = \{v(j, L_j + n_c) < A\}$, $B_3 = \{V(j, L_j + n_c)/v(j, L_j + n_c) \ge 1/2\}$, and $B_4 = \{\sum_{i=L_j+1}^{L_j+n_c} Z_i \le -8\log(c)\}$. Because $m_j = h_j \wedge (L_j + n_c) \wedge L_{j+1}$, we will typically see $m_j = L_j + n_c$. Note that $P_\infty\{\overline{B}_3 \cup \overline{B}_4 \cup B_3B_4\} = 1$. From Lemma 2 and the definition of n_c,

$$\begin{aligned}
&E_\infty\{v(j, m_j);\, v(j, m_j) < A \mid \mathscr{F}_{L_{j-1}}\}\\
&\quad \le A\frac{n_c}{A/c} + E_\infty\{2V(j, L_j + n_c);\, B_1B_3B_4 \mid \mathscr{F}_{L_{j-1}}\}\\
&\qquad + E_\infty\{v(j, L_j + n_c);\, B_2\overline{B}_4 \mid \mathscr{F}_{L_{j-1}}\} + E_\infty\{v(j, L_j + n_c);\, B_2\overline{B}_3 \mid \mathscr{F}_{L_{j-1}}\}\\
&\quad \le cn_c + 2Ac^{-8} + Ac^{-4} + A\,P_\infty\{\overline{B}_3 \mid \mathscr{F}_{L_{j-1}}\}.
\end{aligned}$$

The lemma follows by direct application of Lemma 6. □

LEMMA 13. *Let* $j > 1$ *and* $\xi > 0$. *There exist constant* $k^{(13)}$ *and positive functions* $c^{(13)}(\xi)$, $b^{(13)}(c, \xi)$ *and* $A^{(13)}(c, \xi)$ *such that, on* $\{L_1 > A/c^4\}$,

$$\begin{aligned}
&\Big| E_\infty\{v(j, m_j);\, v(j, m_j) \ge A \mid \mathscr{F}_{L_{j-1}}\} - A\Delta\, P_\infty\{v(j, m_j) \ge A \mid \mathscr{F}_{L_{j-1}}\} \Big|\\
&\qquad \le k^{(13)}A\big[\xi\, P_\infty\{v(j, m_j) \ge A \mid \mathscr{F}_{L_{j-1}}\} + D_{j-1}\big],
\end{aligned}$$

where $D_{j-1} \in \mathscr{F}_{L_{j-1}}$ *is a Bernoulli variate with* $E_\infty\{D_{j-1}\} < \exp(-b^{(13)}A)$ *whenever* $c > c^{(13)}(\xi)$ *and* $A > A^{(13)}(c, \xi)$.

PROOF. Assume without loss that $\xi < 1$. Because $v(j, L_j + n)$ is a martingale,

$$\begin{aligned}
&E_\infty\{v(j, m_j);\, v(j, m_j) \ge A \mid \mathscr{F}_{L_{j-1}}\}\\
&\qquad = E_\infty\{Q(j, L_j) \mid \mathscr{F}_{L_{j-1}}\} - E_\infty\{v(j, m_j);\, v(j, m_j) < A \mid \mathscr{F}_{L_{j-1}}\}.
\end{aligned}$$

We now bound the last term. We use Lemma 12 and the martingale property of $v(j, L_j + n)$ to obtain for A sufficiently large the bounds

$$\begin{aligned} &E_\infty\{Q(j, L_j) \mid \mathscr{F}_{L_{j-1}}\} - (cn_c + 3Ac^{-4} + AD_{j-1}^{(12)}) \\ &\quad \le E_\infty\{v(j, m_j);\, v(j, m_j) \ge A \mid \mathscr{F}_{L_{j-1}}\} \\ &\quad \le E_\infty\{Q(j, L_j) \mid \mathscr{F}_{L_{j-1}}\} \le \kappa(1)\frac{A}{c} \end{aligned} \tag{20}$$

on the event $\{L_1 > A/c^4\}$; the last inequality follows from Condition 1C.

By applying Condition 1C to $E_\infty\{Q(j, L_j) \mid \mathscr{F}_{L_{j-1}}\}$, we may for $\varepsilon \in (0,1)$ use Lemma 11 with (20) to find $c^{**}(\varepsilon)$, $b^{**}(c, \varepsilon)$ and $A^{**}(c, \varepsilon)$ for which

$$\begin{aligned} (1-\varepsilon)E_\infty\{Q(j, L_j) \mid \mathscr{F}_{L_{j-1}}\} - AD_{j-1}^{(12)} &\le E_\infty\{v(j, m_j);\, v(j, m_j) \ge A \mid \mathscr{F}_{L_{j-1}}\} \\ &\le E_\infty\{Q(j, L_j) \mid \mathscr{F}_{L_{j-1}}\} \end{aligned}$$

and

$$\begin{aligned} &(1-\varepsilon)\frac{E_\infty\{Q(j, L_j) \mid \mathscr{F}_{L_{j-1}}\}}{A\Delta} - k^{(11)}D_{j-1}^{(11)} \\ &\quad \le P_\infty\{v(j, m_j) \ge A \mid \mathscr{F}_{L_{j-1}}\} \\ &\quad \le (1+\varepsilon)\frac{E_\infty\{Q(j, L_j) \mid \mathscr{F}_{L_{j-1}}\}}{A\Delta} + k^{(11)}D_{j-1}^{(11)} \end{aligned}$$

both hold on $\{L_1 > A/c^4\}$ whenever $c > c^{**}(\varepsilon)$ and $A > A^{**}(c, \varepsilon)$. Now take the difference of the latter two inequalities, with appropriate choice of ε. □

LEMMA 14. *There exist functions $c^{(14)}(\xi)$ and $A^{(14)}(c, \xi)$ such that, given $0 < \xi < 1$,*

$$\left| \frac{E_\infty\{v(J^*, m_{J^*})\}}{A\Delta} - 1 \right| \le \xi,$$

whenever $c > c^{(14)}(\xi)$ and $A > A^{(14)}(c, \xi)$.

PROOF. Let B_0 denote the event $\{L_1 < A/c^4\}$. Denote by B_j the event $\{v(j, m_j) < A\}$, and let E_j denote the event $\{\sum_{i=2}^{j} D_{i-1}^{(13)} = 0\}$. Recall that $m_j = L_{j+1} \wedge (L_j + n_c) \wedge h_j$. Because $m_j \notin \mathscr{F}_{L_j}$, care is needed when applying our previous lemmas. Also note that $E_j \in \mathscr{F}_{L_{j-1}}$.

We decompose the expected truncated time to exceedance as

$$E_\infty\{v(J^*, m_{J^*})\} = E_\infty\{v(J^*, m_{J^*});\, \overline{E}_{J^*}\} \tag{21}$$

$$\qquad + E_\infty\{v(J^*, m_{J^*});\, E_{J^*}B_0\} \tag{22}$$

$$\qquad + E_\infty\left\{v(c^3+1, m_{c^3+1});\, E_{c^3+1}\overline{B}_0 \bigcap_{i=1}^{c^3+1} B_i\right\} \tag{23}$$

$$+ E_\infty\{v(1, m_1);\ \overline{B}_0\overline{B}_1\} \tag{24}$$

$$+ \sum_{j=2}^{c^3+1} E_\infty\left\{v(j, m_j);\ E_j\overline{B}_0\overline{B}_j \bigcap_{i=1}^{j-2} B_i\right\} \tag{25}$$

$$- \sum_{j=2}^{c^3+1} E_\infty\left\{v(j, m_j);\ E_j\overline{B}_0\overline{B}_j\overline{B}_{j-1} \bigcap_{i=1}^{j-2} B_i\right\}, \tag{26}$$

where intersection over an empty set of indices is the sample space.

Because $v(j, L_j+n)$ is a martingale, we may use Condition 1C and Lemma 2 to show (22) and (24) can be made small by by taking c large. Use Lemma 12 to show (23) is negligible for A large enough.

We next bound the absolute value of (26) by

$$\begin{aligned}\sum_{j=2}^{c^3+1} & E_\infty\left(E_\infty\left\{v(j, m_j);\ \overline{B}_{j-1} \bigcap_{i=1}^{j-2} B_i \,\Big|\, \mathscr{F}_{L_j}\right\}\right) \\ &= \sum_{j=2}^{c^3+1} E_\infty\left\{[Q(j, L_j) - Q(j, m_{j-1})] + Q(j, m_{j-1});\ \overline{B}_{j-1} \bigcap_{i=1}^{j-2} B_i\right\}.\end{aligned} \tag{27}$$

Decompose $Q(j, L_j) - Q(j, m_{j-1}) = \sum_{k=m_{j-1}+1}^{L_j} \Lambda_k^{L_j} + \sum_{k=L_{j-1}+1}^{m_{j-1}} (\Lambda_k^{L_j} - \Lambda_k^{m_{j-1}})$, for $j > 1$. Because $v(j, L_j + n)$ is a martingale and because $L_j \geq m_{j-1}$, the second sum in the decomposition has conditional expectation 0 given $\mathscr{F}_{m_{j-1}}$. Consider now $L_j^* = \min\{n \mid \sum_{k=m_{j-1}+1}^{n} \Lambda_k^n \geq A/c\}$. Because $L_j \leq L_j^*$, and from Condition 1C, the first sum in the decomposition has conditional expectation given $\mathscr{F}_{m_{j-1}}$ bounded by $\kappa(1)A/c$.

Because the events $\overline{B}_{j-1} \bigcap_{i=1}^{j-2} B_i \in \mathscr{F}_{m_{j-1}}$ are disjoint, we may bound (27)—and hence the magnitude of (26)—by

$$\frac{\kappa(1)A}{c} + E_\infty\left\{\sum_{j=2}^{c^3+1} \sum_{\mu=L_{j-1}+1}^{L_{j-1}+n_c} \sum_{k=L_{j-1}+1}^{L_{j-1}+n_c} \Lambda_k^\mu\right\} \leq \frac{\kappa(1)A}{c} + c^3 n_c^2 = \hat{o}_c(A). \tag{28}$$

Note for use in the future that in proving (28) we have obtained a bound for $\sum_{j=2}^{c^3+1} E_\infty\{Q(j, L_j);\ \overline{B}_{j-1} \bigcap_{i=1}^{j-2} B_i\}$. Therefore the Markov inequality yields

$$\sum_{j=2}^{c^3+1} P_\infty\left\{\overline{B}_j\overline{B}_{j-1} \bigcap_{i=1}^{j-2} B_i\right\} \leq \frac{\kappa(1)}{c} + \frac{c^3 n_c^2}{A} = \hat{o}_c(1). \tag{29}$$

Note that $v(J^*, m_{J^*}) \leq \sum_{j=1}^{c^3+1} v(j, m_j)$. Because $\overline{E}_j \in \mathscr{F}_{j-1}$ we use Condition 1C and Lemma 13 to show (21) is small. We next deal with (25). Note that $E_j\overline{B}_0 \bigcap_{i=1}^{j-2} B_i$ are in $\mathscr{F}_{L_{j-1}}$ when $j > 1$. We may use Lemma 13 to show that for c and A large enough, (25) lies between

$$A\Delta\left(1 \pm \frac{\xi}{2}\right)\left[\sum_{j=2}^{c^3+1} P_\infty\left\{E_j\overline{B}_0\overline{B}_j \bigcap_{i=1}^{j-2} B_i\right\}\right]. \tag{30}$$

We show the sum of probabilities in (30) is close to 1. Note that

$$\sum_{j=2}^{c^3+1} P_\infty\left\{E_j\overline{B}_0\overline{B}_j\bigcap_{i=1}^{j-2}B_i\right\} = \sum_{j=2}^{c^3+1} P_\infty\left\{\overline{B}_0\overline{B}_j\bigcap_{i=1}^{j-2}B_i\right\} - \sum_{j=2}^{c^3+1} P_\infty\left\{\overline{E}_j\overline{B}_0\overline{B}_j\bigcap_{i=1}^{j-2}B_i\right\}, \tag{31}$$

and that the second sum in (31) is bounded by $c^6e^{-b^{(13)}A}$. We reexpress the first sum in (31) as

$$P_\infty\left\{\overline{B}_0 \text{ and } J \le c^3+1\right\} - P_\infty\left\{\overline{B}_0\overline{B}_1\right\} + \sum_{j=2}^{c^3+1} P_\infty\left\{\overline{B}_0\overline{B}_j\overline{B}_{j-1}\bigcap_{i=1}^{j-2}B_i\right\}. \tag{32}$$

Now use Lemma 11, Lemma 2, Condition 1C and (29) to conclude that (32) $= 1 + \hat{o}_c(1)$ so that (25) is contained in the interval determined by $A\Delta[1 \pm (\xi/2) \pm \hat{o}_c(1)]$. Combining the bounds completes the proof. □

LEMMA 15. *There exist constants $k^{(15)}$ and $c^{(15)}$, and a function $A^{(15)}(c)$ such that*

$$E_\infty\{v(J_Y^*, m_{J_Y^*});\ Y < m_{J^*}\} \le k^{(15)}A\bigl(P_\infty\{Y < m_{J^*}\} + c^{-1}\bigr),$$

for any stopping time Y, whenever $c > c^{(15)}$ and $A > A^{(15)}(c)$.

PROOF. Define $\underline{J} = \max\{j|L_j \le Y\}$, indexing the last excursion over A/c before Y. We decompose

$$\begin{aligned} v(J_Y^*, m_{J_Y^*}) &\le \sum_{j=1}^{\underline{J}-1} v(j, m_j) + v(\underline{J}, m_{\underline{J}}) \\ &\quad + v(\underline{J}+1, m_{\underline{J}+1}) + v(J_Y^*, m_{J_Y^*})I_{\{J_Y^* \ge \underline{J}+2\}}. \end{aligned} \tag{33}$$

We deal with the contributions to (33) in order. By definition, $L_j \le m_j \le L_{j+1}$. Note $E_\infty\{v(j, m_j);\ L_{j+1} \le Y < m_{J^*}\} \le E_\infty\{v(j, m_j);\ m_j < m_{J^*}\}$ for $j < \underline{J}$. On $\{m_j < m_{J^*}\}$, we have $v(j, m_j) < A$. Now use Condition 1C to bound $E_\infty\{v(1, m_1)\} = E_\infty\{v(1, L_1)\}$ and Lemmas 6, 2 and the definition of n_c for $2 \le j \le c^3$ to show for c large and fixed and A large enough that

$$\sum_{j=1}^{c^3} E_\infty\{v(j, m_j);\ m_j < m_{J^*}\} < 4\kappa(1)A/c.$$

On $\{m_{\underline{J}} \le Y < m_{J^*}\}$, we must have $v(\underline{J}, m_{\underline{J}}) < A$, for otherwise $m_{\underline{J}} = m_{J^*}$. To bound the second piece, observe that $v(\underline{J}, n \vee Y)$ is a martingale and that $v(\underline{J}, Y) < A$ when $Y < m_{\underline{J}}$. Break $\{Y < m_{J^*}\}$ into $\{m_{\underline{J}} \le Y < m_{J^*}\} \cup \{Y < m_{J^*} \wedge m_{\underline{J}}\}$ and integrate to obtain

$$E_\infty\{v(\underline{J}, m_{\underline{J}});\ Y < m_{J^*}\} \le A\,P_\infty\{Y < m_{J^*}\}. \tag{34}$$

Consider now the third piece of (33). Note that $\underline{J}+1$ indexes the first L_j after time Y, so $L_{\underline{J}+1}$ is a stopping time. Also, $v(\underline{J}+1, n \vee L_{\underline{J}+1})$ and $\sum_{k=L_{\underline{J}}+1}^{Y} \Lambda_k^{Y+n}$ are martingales and $\{Y < m_{J^*}\} \in \mathscr{F}_Y \subset \mathscr{F}_{L_{\underline{J}+1}}$. Hence,

$$E_\infty\{v(\underline{J}+1, m_{\underline{J}+1});\ Y < m_{J^*}\}$$
$$= E_\infty\left\{\sum_{k=L_{\underline{J}}+1}^{Y} \Lambda_k^Y;\ Y < m_{J^*}\right\} + E_\infty\left\{\sum_{k=Y+1}^{L_{\underline{J}+1}} \Lambda_k^{L_{\underline{J}+1}};\ Y < m_{J^*}\right\},$$

Now use $Y < L_{\underline{J}+1}$ to bound pointwise the sum in the first expectation, and use the stopping time $\inf\{n \mid \sum_{k=Y+1}^{n} \Lambda_k^n \geq A/c\} \geq L_{\underline{J}+1}$ with Condition 1C to bound the second.

We finally bound the last piece of (33). Let $B_j = \{v(j, m_j) < A\}$ and let $B_0 = \{L_1 < A/c^4\}$. Use Lemma 13 and (34) to show that for c and A sufficiently large,

$$E_\infty\{v(J_Y^*, m_{J_Y^*}) I_{\{J_Y^* \geq \underline{J}+2\}};\ Y < m_{J^*}\}$$
$$\leq 2A\Delta \sum_{j=2}^{c^3+1} P_\infty\left\{\overline{B}_0 \overline{B}_j \bigcap_{i=1}^{j-2} B_i \text{ and } Y < m_{J^*}\right\} + c^3 k^{(13)} A \exp(-b^{(13)} A) \tag{35}$$
$$+ \sum_{j=2}^{c^3+1} E_\infty\{v(j, m_j);\ B_0\} + A\, P_\infty\{Y < m_{J^*}\}.$$

Apply Condition 1C and then Lemma 2 to obtain for $j \geq 2$,

$$E_\infty\{v(j, m_j);\ B_0\} = E_\infty\{Q(j, L_j);\ B_0\} \leq \kappa(1)\frac{A}{c}\, P_\infty\{B_0\} \leq \frac{\kappa(1)}{c^4} A.$$

Because both the odd and even indexed sets of (35) are collections of disjoint events, we obtain for A large a bound of desired form. □

4.3. *Proof of Theorem* 1. Recall the stopping times L_j, with associated continuation processes $v(j, \cdot)$, and auxiliary quantities such as J, are defined in terms of target A and integer c. We now introduce a constant $\varepsilon > 0$. For convenience, we always choose $\varepsilon < 1$ rational and then choose c to make $\pm\varepsilon c$ an integer.

We require three sets of random times, one set corresponding to A and c, and two other sets corresponding to choice of $(1 \pm \varepsilon)A$ and $(1 \pm \varepsilon)c$. Note that the stopping times L_j are the same, regardless of the pair specified. Therefore the continuation processes $v(j, \cdot)$ and $V(j, \cdot)$ are also the same under any specification. Where the choice makes a difference, we use the subscript $\pm\varepsilon$. For example, $h_{-\varepsilon, j}$ denotes the first time n that $v(j, n) \geq (1-\varepsilon)A$. Similarly, $m_{+\varepsilon, j}$ is defined to be $h_{+\varepsilon, j} \wedge (L_j + n_{(1+\varepsilon)c}) \wedge L_{j+1}$.

We show first that $E_\infty\{R_{N_A}\} = E_\infty\{N_A\}$. From Lemma 1, the latter is finite. For n large, $E_\infty\{R_{N_A\wedge n} - N_A \wedge n\} = 0$ from the martingale property. Equality follows from monotone convergence and Fatou's lemma. Hence it suffices to study $E_\infty\{R_{N_A}\}$.

We now argue that $\liminf_{A\to\infty} E_\infty\{R_{N_A}\}/A \ge \Delta$. Let ξ, $\varepsilon < 1/2$, and c be positive, whose exact values will be chosen and fixed in that order before we let $A \to \infty$. Given ξ and ε, use Lemma 14 to show

$$1 - \xi < \frac{E_\infty\{v(J^*_{\pm\varepsilon}, m_{\pm\varepsilon,\, J^*_{\pm\varepsilon}})\}}{(1 \pm \varepsilon)A\Delta} < 1 + \xi \tag{36}$$

for $c > 2c^{(14)}(\xi)$ and $A > 2A^{(14)}((1 \pm \varepsilon)c, \xi)$. With $v_{-\varepsilon} = v(J^*_{-\varepsilon}, m_{-\varepsilon, J^*_{-\varepsilon}})$, we obtain

$$\begin{aligned} E_\infty\{R_{N_A}\} &\ge E_\infty\big\{R_{N_A};\ N_A \ge m_{-\varepsilon,\, J^*_{-\varepsilon}}\big\} \\ &\ge E_\infty\{v_{-\varepsilon}\} - E_\infty\big\{v_{-\varepsilon};\ N_A < m_{-\varepsilon,\, J^*_{-\varepsilon}}\big\}. \end{aligned} \tag{37}$$

Define events

$$\begin{aligned} B_1 &= \{N_A - L_W > n_c\}, \\ B_2 &= \{W \ge c^3\}, \\ B_3 &= \Big\{\min_{2\le j\le c^3} L_j - L_{j-1} \le 2n_c\Big\}, \\ B_4 &= \Big\{\max_{2\le j\le c^3} \sup_{n\ge 2n_c} v(j, L_j + n) > A/c^9\Big\}, \\ B_5 &= \Big\{\sup_{n\ge L_W} v(1, n) > \varepsilon A/4\Big\}. \end{aligned}$$

Consider now $\{N_A < m_{-\varepsilon,\, J^*_{-\varepsilon}}\} \subset \{N_A < m_{-\varepsilon,\, J^*_{-\varepsilon}}$ and $\bigcap_{i=1}^4 \overline{B}_i\} \cup \bigcup_{i=1}^4 B_i$. We know $Q(W, N_A) < A/c$, for otherwise $L_{W+1} \le N_A$, contradicting the definition of W. On the event $\bigcap_{i=1}^4 \overline{B}_i$,

$$\begin{aligned} R_{N_A} &= Q(W, N_A) + v(W, N_A) + \sum_{j=1}^{W-1} v(j, N_A) \\ &< A/c + v(W, N_A) + c^3 A/c^9 + \sup_{n\ge L_W} v(1, n). \end{aligned}$$

Hence $v(W, N_A)/A \ge v(W, N_A)/R_{N_A} > 1 - 2/c - \varepsilon/4$ on $\bigcap_{i=1}^5 \overline{B}_i$. From Lemmas 5, 7 and 8, the probability of $\bigcup_{i=1}^4 B_i$ is less than $(1 + k^{(7)} + 2k^{(8)})/c$, for c and A sufficiently large. Also,

$$P_\infty\{B_5\} = E_\infty P_\infty\left\{\sup_{n\ge L_W} v(1, n) > \frac{\varepsilon A}{4} \,\Big|\, \mathscr{F}_{L_1}\right\} \le E_\infty\left\{\frac{v(1, L_1)}{\varepsilon A/4}\right\} \le \frac{4\kappa(1)}{\varepsilon c}.$$

When c is chosen to make $\varepsilon > 4/c$, we have $N_A \geq m_{-\varepsilon,\, J^*_{-\varepsilon}}$ on $\bigcap_{i=1}^5 \overline{B}_i$, so that

$$P_\infty\left\{N_A < m_{-\varepsilon,\, J^*_{-\varepsilon}}\right\} < \left(1 + k^{(7)} + 2k^{(8)} + 4\kappa(1)\varepsilon^{-1}\right)/c. \tag{38}$$

Note both $m_{-\varepsilon,\, J^*_{-\varepsilon}}$ and $m_{-\varepsilon,\, J^*_{-\varepsilon, N_A}}$ are stopping times and $m_{-\varepsilon,\, J^*_{-\varepsilon, N_A}} \leq m_{-\varepsilon,\, J^*_{-\varepsilon}}$ on $\{N_A < m_{-\varepsilon,\, J^*_{-\varepsilon}}\}$. Hence, given fixed ξ and ε, use (36), (37), (38) with Lemma 15 to choose c sufficiently large, so that

$$\begin{aligned} E_\infty\{R_{N_A}\} &\geq E_\infty\{v_{-\varepsilon}\} - E_\infty\left\{v_{-\varepsilon};\ N_A < m_{-\varepsilon,\, J^*_{-\varepsilon}}\right\} \\ &= E_\infty\{v_{-\varepsilon}\} - E_\infty\left\{v(J^*_{-\varepsilon, N_A}, m_{-\varepsilon,\, J^*_{-\varepsilon, N_A}});\ N_A < m_{-\varepsilon,\, J^*_{-\varepsilon}}\right\} \\ &\geq (1-\xi)(1-\varepsilon)A\Delta - A\varepsilon \end{aligned}$$

for all A sufficiently large. Hence $\liminf_{A\to\infty} E_\infty\{R_{N_A}\}/A \geq \Delta$.

Finally, we show $\limsup_{A\to\infty} E_\infty\{R_{N_A}\}/A \leq \Delta$, completing the proof of Theorem 1. Note that $N_A \leq m_{+\varepsilon,\, J_{+\varepsilon}}$ for any $\varepsilon > 0$. Define and redefine

$$\begin{aligned} B_0 &= \left\{L_1 \leq A/c^4\right\} \\ B_1 &= \left\{N_A \geq m_{+\varepsilon,\, J^*_{+\varepsilon}}\right\} \\ B_2 &= \left\{J_{+\varepsilon} > (1+\varepsilon)^3 c^3\right\}. \end{aligned}$$

Write $m_+ = m_{+\varepsilon,\, J^*_{+\varepsilon}}$. Because $N_A = m_+$ on $B_1\overline{B}_2$, we obtain

$$\begin{aligned} E_\infty\{R_{N_A}\} &= E_\infty\left\{R_{m_+}\right\} + E_\infty\left\{(R_{N_A} - R_{m_+});\ B_1B_2 \cup \overline{B}_1\right\} \\ &\leq E_\infty\left\{R_{m_+}\right\} - E_\infty\left\{(R_{m_+} - R_{N_A});\ \overline{B}_1\right\} + E_\infty\left\{R_{N_A};\ B_1B_2\right\}. \end{aligned} \tag{39}$$

We deal with the contributions to (39) in order. Write $c_\varepsilon = (1+\varepsilon)c$. We decompose R_{m+} into its exceedance epochs to obtain

$$R_{m_+} = Q(J^*_{+\varepsilon} + 1, m_+) + v(J^*_{+\varepsilon}, m_+) + \sum_{j=1}^{c_\varepsilon^3} v(j, m_+) I_{\{m_j < m_+\}}.$$

The first term is bounded by $\sum_{j=1}^{c^3+1}\sum_{l=1}^{n_{c_\varepsilon}} Q(j, L_j + l)$ which accounts for $(c_\varepsilon^3 + 1)n_{c_\varepsilon}(n_{c_\varepsilon} + 1)/2$ likelihood ratios, each having expectation 1. Apply (36) to the second term. Hence $E_\infty\{R_{m_+}\}$ is bounded above by

$$\begin{aligned} &\left[(c_\varepsilon^3 + 1)\frac{n_{c_\varepsilon}(n_{c_\varepsilon} + 1)}{2}\right] + (1+\xi)(1+\varepsilon)A\Delta + E_\infty\{v(1, m_+)\} \\ &\quad + \sum_{j=2}^{c_\varepsilon^3} E_\infty\left\{v(j, m_+);\ \overline{B}_0 \cap \{m_j < m_+\}\right\} + \sum_{j=2}^{c_\varepsilon^3} E_\infty\{v(j, m_+);\ B_0\}. \end{aligned}$$

Because $v(j, L_j + n)$ is a martingale and $v(j, L_j) = Q(j, L_j)$, we can apply, respectively, Condition 1C, Lemma 12 and Lemma 2 to obtain

$$
\begin{aligned}
E_\infty\{R_{m_+}\} < &\left[(c_\varepsilon^3 + 1)\frac{n_{c_\varepsilon}(n_{c_\varepsilon}+1)}{2}\right] + (1+\xi)(1+\varepsilon)A\Delta \\
&+ \kappa(1)\frac{A}{c} + \left[c_\varepsilon^4 n_{c_\varepsilon} + \frac{3A}{c} + (1+\varepsilon)A\exp(-b^{(12)}A)\right] \\
&+ c_\varepsilon^3 \kappa(1)\frac{A}{c^4},
\end{aligned}
\tag{40}
$$

bounding the first contribution to (39).

Note N_A and m_+ are both stopping times, and $\overline{B}_1 \in \mathscr{F}_{N_A \wedge m_+}$, hence

$$
E_\infty\{(R_{m_+} - R_{N_A}); \overline{B}_1\} = E_\infty\{(m_+ - N_A); \overline{B}_1\} > 0,
$$

by the martingale property, so we can ignore the second piece of (39).

We next bound the third contribution to (39). On $B_1 B_2$, we have $Q(c_\varepsilon^3 + 1, N_A) \le Q(c_\varepsilon^3 + 1, N^*)$ for $N^* = \inf_{n \ge L_{c_\varepsilon^3+1}} \{n \mid Q(c_\varepsilon^3+1, n) \ge A\}$. Now take expectations, applying Condition 1C, Lemmas 2, 8, 12 and the martingale property of the $v(j, \cdot)$, to obtain

$$
\begin{aligned}
E_\infty\{R_{N_A}; B_1 B_2\} \le &\ \kappa(1)A\, P_\infty\{B_2\} + \kappa(1)\frac{A}{c} \\
&+ \sum_{j=2}^{c_\varepsilon^3}\left(c_\varepsilon n_{c_\varepsilon} + \frac{3(1+\varepsilon)A}{c_\varepsilon^4} + (1+\varepsilon)A\exp(-b^{(12)}A)\right) \\
&+ c_\varepsilon^3\kappa(1)\frac{A}{c}\, P_\infty\{B_0\}.
\end{aligned}
\tag{41}
$$

Finally, add (40) and (41), choose c large and then let $A \to \infty$. □

REFERENCES

GORDON, L. and POLLAK, M. (1994). An efficient sequential nonparametric scheme for detecting a change of distribution. *Ann. Statist.* **22** 763–804.

GORDON, L. and POLLAK, M. (1995). A robust surveillance scheme for stochastically ordered alternatives. *Ann. Statist.* **23** 1350–1375.

LORDEN, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Statist.* **42** 1897–1908.

LORDEN, G. and EISENBERGER, I. (1973). Detection of failure rate increases. *Technometrics* **15** 167–175.

MOUSTAKIDES, G. V. (1986). Optimal stopping times for detecting changes in distributions. *Ann. Statist.* **14** 1379–1387.

PAGE, E. S. (1954). Continuous inspection schemes. *Biometrika* **41** 100–115.

POLLAK, M. (1985). Optimal detection of a change in distribution. *Ann. Statist.* **13** 206–227.

POLLAK, M. (1987). Average run lengths of an optimal method of detecting a change in distribution. *Ann. Statist.* **15** 749–779.

POLLAK, M. and SIEGMUND, D. (1991). Sequential detection of a change in a normal mean when the initial value is unknown. *Ann. Statist.* **19** 394–416.

RITOV, Y. (1990). Decision theoretic optimality of the CUSUM procedure. *Ann. Statist.* **18** 1464–1469.

ROBERTS, S. W. (1966). A comparison of some control chart procedures. *Technometrics* **8** 411–430.

SHEWHART, W. A. (1931). *The Economic Control of the Quality of Manufactured Product*. Macmillan, New York.

SHIRYAYEV, A. N. (1963). On optimum methods in quickest detection problems. *Theory Probab. Appl.* **8** 22–46.

SIEGMUND, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*. Springer, New York.

VAN DOBBEN DE BRUYN, C. S. (1968). *Cumulative Sum Tests: Theory and Practice*. Hafner, New York.

WILSON, D. W., GRIFFITHS, K., KEMP, V. W., NIX, A. B. J. and ROWLANDS, R. J. (1979). Interval quality control of radioimmunoassays: monitoring of error. *Journal of Endocrinology* **80** 365–372.

YAKIR, B. (1997). Optimal detection of a change in distribution when the observations are independent. *Ann. Statist.* To appear.

FILOLI INFORMATION SYSTEMS
777 CALIFORNIA AVENUE
PALO ALTO, CA 94304

DEPARTMENT OF STATISTICS
HEBREW UNIVERSITY
JERUSALEM 91905
ISRAEL