Estimating a Random Walk First-Passage Time from Noisy or Delayed Observations

Marat V. Burnashev and Aslan Tchamkerten, Member, IEEE

Abstract—A random walk (or a Wiener process), possibly with drift, is observed in a noisy or delayed fashion. The problem considered in this paper is to estimate the first time \( t \) the random walk reaches a given level. Specifically, the \( p \)-moment \( (p \geq 1) \) optimization problem \( \inf_{\eta} \mathbb{E}|\eta - t|^p \) is investigated where the infimum is taken over the set of stopping times that are defined on the observation process.

When there is no drift, optimal stopping rules are characterized for both types of observations. When there is a drift, upper and lower bounds on \( \inf_{\eta} \mathbb{E}|\eta - t|^p \) are established for both types of observations. The bounds are tight in the large-level regime for noisy observations and in the large-level-large-delay regime for delayed observations. Noteworthy, for noisy observations there exists an asymptotically optimal stopping rule that is a function of a single observation.

Simulation results are provided that corroborate the validity of the results for non-asymptotic settings.

Index Terms—change-point detection problem, estimation, optimal stopping theory, random walk, stopping time, tracking stopping time (TST), Wiener process

I. INTRODUCTION

Suppose \( X = \{X_t\}_{t \geq 0} \) is a stochastic process and \( \tau \) a stopping time defined over \( X \). Statistician has access to \( X \) only through correlated observations \( Y = \{Y_t\}_{t \geq 0} \) and wishes to find a stopping \( \eta \) defined over \( Y \) that gets as close as possible to \( \tau \), for instance, so as to minimize some average absolute moment \( \mathbb{E}|\eta - \tau|^p \). This general formulation was introduced in [9] as the Tracking Stopping Time (TST) problem, and an early instance of it where \( \eta = \tau \) was investigated in [8].

The TST problem generalizes the long studied Bayesian change-point detection problem (see, e.g., [13] and the books [10] and [1] for surveys on theory and applications of the change-point problem).

In the Bayesian change-point problem, there is a random variable \( \theta \), taking on values in the positive integers, and two probability distributions \( P_0 \), the “nominal” distributions, and \( P_1 \), the “alternative” distribution. Under \( P_0 \), the conditional density function of \( Y_t \) given \( Y_0, Y_1, \ldots, Y_{t-1} \) is \( f_0(Y_t|Y_0, Y_1, \ldots, Y_{t-1}) \), for every \( t \geq 0 \). Under \( P_1 \), the conditional density function of \( Y_t \) given \( Y_0, Y_1, \ldots, Y_{t-1} \) is \( f_1(Y_t|Y_0, Y_1, \ldots, Y_{t-1}) \), for every \( t \geq 0 \). The observed process is distributed according \( P_0 \), which assigns the conditional density functions of \( P_0 \) for all \( t < \theta \), and the conditional density functions of \( P_1 \) for all \( t \geq \theta \). The Bayesian change-point problem typically consists in finding a stopping time \( \eta \) with respect to \( \{Y_t\} \), that minimizes some (loss) function of the delay \( \eta - \theta \).

To see that the Bayesian change-point problem can always be formulated as a TST problem, it suffices to define the process \( X = \{X_t\}_{t \geq 0} \) as \( X_t = 0 \) for \( t < \theta \) and \( X_t = 1 \) for \( t \geq \theta \). The Bayesian change-point problem becomes the TST problem which consists in tracking \( \theta \) (now defined as a stopping time with respect to \( X \)) through \( Y \).

The difference between the Bayesian change-point problem and the TST problem lies in the equality

\[ \mathbb{P}(\theta = t | X_t > n) = \mathbb{P}(\theta = t | Y_t > n), \quad k > n \]

which always holds for the former but need not hold for the latter [9]. In other words, for TST problems past observations are in general useful for estimating the future value of \( \tau \), by contrast with Bayesian change-point problems. For specific applications of the TST problem formulation related to monitoring, communication, and forecasting we refer to [9] Section I.

In [9], through a computer science approach, a general algorithmic solution is proposed for constructing optimal “trackers” for the cases where \( X \) and \( Y \) are processes defined over finite alphabets and \( \tau \) is bounded. What motivated an algorithmic approach is that the TST problem generalizes the Bayesian change-point problem for which general closed-form analytical solutions have been reported only for specific asymptotic regimes, typically the vanishing false-alarm regime (see, e.g., [6]). Non-asymptotic closed-form solutions have been obtained essentially for i.i.d. cases where, conditioned on the change-point value, observations are independent with common distribution \( P_0 \) and \( P_1 \) before and after the change, respectively (see, e.g., [11], [12]).

Two natural TST settings include the ones where the observation process \( Y \) is a noisy or delayed version of \( X \). In this paper we investigate both situations when \( X \) is a Gaussian random walk (or a Wiener process) possibly with drift, and \( \tau \) is the first time when \( X \) reaches some given level \( \ell \). For noisy and delayed observations, we establish lower bounds on

\[ \inf_{\eta} \mathbb{E}|\eta - \tau|^p, \quad p \geq 1 \]
where the infimum is over all stopping times with respect to Y, then exhibit stopping rules that achieve these bounds in the large-threshold regime and large-delay-large-threshold regime, respectively. For noisy observations, two complementary asymptotically optimal stopping rules are proposed. One depends on a single observation at some fixed time but its optimality is usually very asymptotic. The other performs a sequential minimum mean square error (mmse) estimate of $X_\ell$ given $Y_t$, $t = 0, 1, \ldots$ and stops as soon as this estimate reaches level $\ell$. As such, the second stopping time needs many more observations, roughly $\ell/s$, but performs significantly better in the non-asymptotic regime.

In the particular case where $\tilde{X}$ doesn’t drift, we characterize $\inf_{\eta} \mathbb{E}[|\eta - \tau|^p]$ non-asymptotically for both the noisy and the delayed observation cases.

Section III contains the main results and Section III is devoted to the proofs.

II. RESULTS

Consider the discrete-time process

$$X : X_0 = 0, \quad X_t = \sum_{i=1}^t V_i + st \quad t \geq 1,$$

where $s > 0$ is some known constant, where $V_1, V_2, \ldots$ are i.i.d. $\sim \mathcal{N}(0, 1)$ (zero mean unit variance Gaussian random variables), and consider the first-passage time

$$\tau_\ell = \inf \{ t \geq 0 : X_t \geq \ell \}$$

for some known fixed threshold level $\ell > 0$.

Given sequential observations of a process $Y = \{Y_t\}_{t \geq 0}$ correlated to $X$, we consider the optimization problem

$$\inf_{\eta} \mathbb{E}[|\eta - \tau|^p], \quad \tau \geq 1,$$

(1)

where the infimum is over all stopping times $\eta$ defined with respect to the natural filtration induced by $Y_t$.

The results, presented in the next two subsections, relate to the situations where $Y$ is either a noisy version of $X$, or a delayed version of $X$.

Throughout the paper the following notational conventions are adopted. We use $\eta$ to denote a function of $Y = Y_\infty^\infty$. When $\eta$ has no argument, such as in (1), we mean that $\eta$ is a stopping time with respect to $Y$. Instead, if $\eta$ has an argument, we mean that $\eta$ is a function of its argument which need not be a stopping time with respect to $Y$. For example, $\eta(Y_\alpha^b)$, with $0 \leq \alpha \leq b \leq \infty$, refers to a function of observations $Y_\alpha^b = Y_\alpha, Y_{\alpha+1}, \ldots, Y_b$.

Further, we frequently omit arguments of functions (or estimators) that appear in expressions to be optimized. For instance, instead of

$$\inf_{\eta(Y_\alpha^b)} \mathbb{E}[|\eta(Y_\alpha^b) - \tau|^p],$$

we simply write

$$\inf_{\eta(Y_\alpha^b)} \mathbb{E}[|\eta - \tau|^p].$$

A. Noisy observations

Consider the observation process

$$Y : Y_0 = 0, \quad Y_t = X_t + \varepsilon \sum_{i=1}^t W_i \quad t \geq 1,$$

where $W_1, W_2, \ldots$ are i.i.d. $\sim \mathcal{N}(0, 1)$ and where $\varepsilon \geq 0$ is some known constant. The observation noises $\{W_t\}$ are supposed to be independent of $\{V_t\}$.

Note that if $\ell = 0$ or if $\varepsilon = 0$ (i.e., $X = Y$), (1) is equal to zero by setting $\eta = 0$ and $\eta = \tau_\ell$, respectively.

Interestingly, when $\ell > 0$, $\varepsilon > 0$, and $s = 0$, it turns out that it is impossible to track $\tau_\ell$, even having access to the entire observation process $Y_\ell^\infty$.

**Theorem 1** (Noisy observations, $s = 0$). For $s = 0$, $\varepsilon > 0$, $\ell > 0$, and $p \geq 1/2$, we have

$$\mathbb{E}[|\eta(Y_\infty^\infty) - \tau|^p] = \infty$$

for any estimator $\eta(Y_\infty^\infty)$ of $\tau_\ell$.

We now consider the case $\ell > 0$, $\varepsilon > 0$, and $s > 0$. The next result characterizes (1) in the limit $\ell \to \infty$ and provides two asymptotically optimal stopping rules. One of these rules is non-sequential in the sense that it depends on a single observation.

The sequential stopping rule is defined as

$$\hat{\eta}_\ell^\circ \equiv \inf \{ t \geq 0 : \hat{X}_t \geq \ell \},$$

(2)

where $\hat{X}_0 \equiv 0$ and where

$$\hat{X}_t \equiv \frac{1}{1 + \varepsilon^2} Y_t + \frac{s \varepsilon^2}{1 + \varepsilon^2} t, \quad t \geq 1$$

(3)

is the mmse estimator of $X_t$ given observation $Y_t$.

The non-sequential stopping rule is defined as follows. Let

$$\eta_0^\circ \equiv t^* + \frac{(\ell - \hat{X}_{t^*})_+}{s},$$

(4)

with

$$t^* \equiv \lfloor \ell/s - (\ell/s)^q \rfloor,$$

(5)

for some arbitrary constant $q \in (1/2, 1)$. Notice that $\eta_0^\circ$ is only a function of observation $Y_{t^*}$.

**Theorem 2** (Noisy observations, $s > 0$). Fix $0 < \varepsilon < \infty$, $0 < s < \infty$, and $p \geq 1$. Then, for $\eta = \eta_0^\circ$ or $\eta = \hat{\eta}_\ell^\circ$

$$\mathbb{E}[|\eta - \tau|^p] = (1 + o(1)) \inf_{\eta(Y_\infty^\infty)} \mathbb{E}[|\eta - \tau|^p]$$

$$= (1 + o(1)) C_\ell (\ell, s, \varepsilon, p)$$

(6)

Recall that $\eta(Y_\infty^\infty)$ denotes an arbitrary function of observations $Y_\infty^\infty$ which need not be a stopping time, according to our notational convention of the previous section.

$x_+$ denotes $\max\{0, x\}$ and $[x]$ denotes the integer part of $x$. 

---

3We consider only non-randomized stopping times since this does not induce a loss of optimality with respect to (1) (see, e.g., [4] Chap. 8.5) where randomization is shown to be useless for general statistical decision problems.
as $\ell \to \infty$, where

$$C_1(\ell, s, \varepsilon, p) \overset{\text{def}}{=} \left( \frac{\ell \varepsilon^2}{s^4(1 + \varepsilon^2)} \right)^{p/2} \mathbb{E} |N|^p,$$

and where $N \sim N(0, 1)$.

Since

$$\mathbb{E} |\eta - \tau_\ell|^p \geq \inf_{\eta'} \mathbb{E} |\eta' - \tau_\ell|^p \geq \inf_{\tilde{\eta}(Y_0^\varepsilon)} \mathbb{E} |\tilde{\eta} - \tau_\ell|^p,$$

the first equality in (6) says that both stopping rules $\eta^*_\ell$ and $\eta^*_{\ell}$ do as well as the best non-causal estimators of $\tau_\ell$ with access to the entire observation process $Y$, asymptotically. Moreover, note that asymptotic optimality is universal over $p \geq 1$ for $\eta^*_\ell$ and universal over both $p$ and $\varepsilon$ for $\eta^*_{\ell}$—since the former does not depend on $p$ and the latter depends neither on $p$ nor on $\varepsilon$. For $p = 1$, the optimality of $\eta^*_\ell$ was established in [2] Theorem 2.3.

Since $\eta^*_{\ell}$ does not exploit the dependency between $X$ and $Y$ ($\eta^*_{\ell}$ does not depend on $\varepsilon$), it may be expected that $\eta^*_{\ell}$ performs significantly better that $\eta^*_\ell$ for moderate to low values of $\ell$. In fact, this claim is supported numerically. An illustration is given by Fig. 1 which represents numerical evaluations of

$$\frac{\mathbb{E} |\eta - \tau_\ell|^p}{C_1(\ell, s, \varepsilon, p)}$$

as a function of $\ell$ for $\eta \in \{\eta^*_\ell, \eta^*_{\ell}, \ell/s\}$, with parameters $p = 1, s = 10, \varepsilon = .5$. The parameter $q$ in the definition of $\eta^*_\ell$ is chosen to be equal to .51. The simulation has a precision of $\delta = .1$ for $\eta = \eta^*_\ell$ and $\eta = \eta^*_{\ell}$, and a precision of $\delta = .5$ for $\delta = \ell/s$. By precision we mean that the numerical evaluation of (7) deviates from it by less than $\delta$ with probability at least $1 - \delta$. Simulation details are provided in the appendix.

We observe that, as $\ell \to \infty$, (7) tends to 1 for both $\eta = \eta^*_\ell$ and $\eta = \eta^*_{\ell}$, as predicted by Theorem 2. However, $\eta^*_{\ell}$ performs significantly better than $\eta^*_{\ell}$ in the non-asymptotic regime. For instance, for $\ell \approx 1000$, $\mathbb{E} |\eta^*_\ell - \tau_\ell|$ is roughly a third of $\mathbb{E} |\eta^*_{\ell} - \tau_\ell|$.

More generally, simulation results suggest that $\mathbb{E} |\eta^*_{\ell} - \tau_\ell|^p$ never exceeds $\mathbb{E} |\eta^*_{\ell} - \tau_\ell|^p$, and this for arbitrary $\ell > 0, s > 0, \varepsilon > 0$, and $p \geq 1$.

Moreover, the difference between $\mathbb{E} |\eta^*_{\ell} - \tau_\ell|^p$ and $\mathbb{E} |\eta^*_{\ell} - \tau_\ell|^p$ increases as $\ell$ decreases, and can be very significant for moderate to low values of $\ell$. For instance, for $\ell = 1000, s = 10, \varepsilon = .1$, and $q = .51$, we have

$$(\mathbb{E} |\eta^*_{\ell} - \tau_\ell|)/\mathbb{E} |\eta^*_{\ell} - \tau_\ell| \approx 12 (!)$$

Thus, $\eta^*_\ell$ is suitable for very large values of $\ell$ since it has the interesting feature of being a function of a single observation. While also asymptotically optimal, $\eta^*_{\ell}$ does significantly better than $\eta^*_\ell$ in the non-asymptotic regime, but requires roughly $\ell/s$ observations on average. To see this, note that $\mathbb{E} X_\ell^\varepsilon \approx \ell$, and since $X_\ell - X_{\ell-1} = s$, we have $\mathbb{E} |\eta^*_{\ell}| \approx \ell/s$ by Wald’s equality—the approximations become equalities if we ignore excess over the boundary (variously known as “overshoot”), i.e., that $X_\ell^\varepsilon$ may exceed $\ell$.

Concerning the fixed time estimator $\eta = \ell/s$, later it is shown (see paragraph after Lemma 1) that

$$\lim_{\ell \to \infty} \frac{\mathbb{E} |\tau_\ell - \ell/s|^p}{C_1(\ell, s, \varepsilon, p)} = \left( \frac{1 + s^2 \varepsilon^p}{s^2} \right)^{p/2}$$

which is always greater than 1. Hence $\eta = \ell/s$ is always suboptimal, and in particular for small values of the noise parameter $\varepsilon$. As $\varepsilon$ increases, the observation process $Y$ becomes noisier and ultimately useless in the limit $\varepsilon \to \infty$. In this regime the fixed time estimator $\ell/s$ is optimal. In the example of Fig. 1 the right-hand side of (7) is equal to $\sqrt{5}$.

**B. Delayed observations**

Consider the observation process

$$Y_i = 0, Y_1 = 0, \ldots, Y_d = 0 \quad Y_i = X_{\ell-d}, \quad t \geq d + 1$$

for some fixed positive integer $d \geq 0$.

Given $d \geq 0, \ell \geq 0$, and $s \geq 0$, define the stopping rule

$$\eta_d \overset{\text{def}}{=} \inf \{t \geq 0 : Y_{t} \geq \ell - s \cdot d \}.$$

Notice that $\eta_d$ is a very natural candidate for estimating $\tau_\ell$ since, on average, $X_t$ is $s \cdot d$ higher than $Y_t$. In fact, the following two theorems establish optimality of $\eta_d$ for any $s \geq 0$.

**Theorem 3** (Delayed observations, $s = 0$). For $s = 0$, $\ell > 0$, and $p \geq 1/2$,

$$\inf_{\eta} \mathbb{E} |\eta - \tau_\ell|^p = d^p = \mathbb{E} |\eta_d - \tau_\ell|^p.$$

Instead, when the drift is positive we have:

**Theorem 4** (Delayed observations, $s > 0$). For $s > 0$ and $p \geq 1$,

$$\inf_{\eta} \mathbb{E} |\eta - \tau_\ell|^p = (1 + o(1))\mathbb{E} |\eta_d - \tau_\ell|^p = (1 + o(1))C_2(d, s, p)$$

as $d \to \infty$ while $\ell = \ell(d) \geq s \cdot d$, where

$$C_2(d, s, p) \overset{\text{def}}{=} \left( \frac{d^{p/2}}{s^p} \right) \mathbb{E} |N|^p.$$
for delayed observations, where
the corresponding estimation error is equal to
is optimal to wait until it is certain that
the estimation error grows as
s > 0
is no more a stopping time with respect to
by their continuous time counterparts; i.e.,
Y_	ext{stoch.} \sim \mathcal{N}(0, \sigma^2)
which evaluates (9) with a finite number of random samples.
The small oscillations around
as a function of
\mu = \inf\{t \geq 1 : S_t \geq \ell\}
Then,
i. the following inequalities hold
\mathbb{P}(\mu < \ell/s - z) \leq \exp \left\{ -\frac{s^2z^2}{2\sigma^2(\ell/s - z)} \right\} \quad (10)
for 0 \leq z < \ell/s;
\mathbb{P}(\mu > \ell/s + z) \leq \exp \left\{ -\frac{s^2z^2}{2\sigma^2(\ell/s + z)} \right\} \quad (11)
for \ z \geq 0;
ii. for any \ p \geq 0
\mathbb{E}\left| \frac{\mu - \ell}{s} \right| \leq k_1(k_2 + \ell)p/2
(12)
where 0 \leq k_1, k_2 < \infty are constants that depend on
p, s, \sigma^2 but not on \ell;
iii. as \ell \to \infty,
\sqrt{s^3(\ell/s - \tau_\ell)} \to \mathcal{N}(0,1)
in distribution.
Claim iii. of Lemma 1 implies 8. To see this, let \tau_\ell be the
first time process \ X \ reaches level \ell. Claim iii. of Lemma 1
then gives
\mathbb{E}|\tau_\ell - \ell/s|^p = (1 + o(1))\ell^{p/2} \sigma^p \mathbb{E}|N|^p \quad (\ell \to \infty) (13)
where \ N \sim \mathcal{N}(0, 1). This establishes (8).
The following basic fact is repeatedly used in the proofs of
Theorems 2 and 4
Fact 1. Let \ (S, Q) be two arbitrary random variables. Then,
\inf_{\eta(S)} \mathbb{E}|\eta f(S) - g(S) - h(Q)|^p = \inf_{\eta(S)} \mathbb{E}|\eta - h(Q)|^p \quad p \geq 0
for any functions \ g(\cdot) and \ h(\cdot), and any function \ f(\cdot) such that \ f(S) > 0 almost surely.
To see this, notice first the obvious inequality
\inf_{\eta(S)} \mathbb{E}|\eta f(S) - g(S) - h(Q)|^p \geq \inf_{\eta(S)} \mathbb{E}|\eta - h(Q)|^p.
To see that
\[ \inf_{\eta(S)} \mathbb{E}|f(S) - g(S) - h(Q)|^p \leq \inf_{\eta(S)} \mathbb{E}|\eta - h(Q)|^p, \]
oberse that for any \( \eta = \eta(S) \) one can find \( \tilde{\eta} = \tilde{\eta}(S) \) such that
\[ \tilde{\eta}f(S) - g(S) = \eta \]
almost surely since \( f(S) > 0 \) almost surely.

To illustrate Fact 1, consider the following simple example, variations of which appear in the proofs of Theorems 2 and 4.

Let \( X = Y + Z \) where \( X \) and \( Y \) are arbitrary random variables. Then, for any \( c > 0 \)
\[ \inf_{\eta(Y)} \mathbb{E}|\eta - c \cdot X|^p = c^p \inf_{\eta(Y)} \mathbb{E}|\eta - c \cdot Y|^p = c^p \inf_{\eta(Y)} \mathbb{E}|\eta - Z|^p, \]
where the last equality follows from Fact 1 with \( S = Y \), \( Q = Z \), \( f(S) = 1/c \), \( g(S) = S \), and \( h(Q) = Q \).

We now prove Theorems 2 and 4, then Theorem 5. Throughout the proofs, \( N \) always denotes a zero mean unit variance Gaussian random variable.

A. Proof of Theorem 2

We first show that
\[ \inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - \tau_t|^p \geq (1 + o(1))C_1(\ell, s, \varepsilon, p), \tag{14} \]
where \( C_1(\ell, s, \varepsilon, p) \) is defined in Theorem 2, then show that \( \mathbb{E}|\eta - \tau_t|^p \) is equal to the right-hand side of (14) for \( \eta = \eta^\circ_t \) and \( \eta = \eta^\circ_t \). Before proceeding formally, we outline the main arguments.

To show (14), the main idea is to reduce the minimization problem of estimating \( \tau_t \) to the one of estimating process \( X \) at an instant close to \( \ell/s \), the expected time \( X \) reaches level \( \ell \). To do this reduction, let \( t^* \) be such that \( t^* \approx \ell/s \) while satisfying \( \mathbb{P}(\tau_t \geq t^*) \approx 1 \)—one such instant is the \( t^* \) defined in (5). It then follows that
\[ \tau_t \overset{d}{=} t^* + \frac{(\ell - \hat{X}_t^* +)}{s}, \tag{15} \]
since the time it takes for \( X \) to go up by \( q \geq 0 \) is \( q/s \) plus some small Gaussian term, by Claim iii. of Lemma 1.

From (15), the fact that \( Y_t^\ast \) is a sufficient statistic for \( X_t^\ast \), and that \( t^* \) is close to \( \ell/s \), one can show that
\[ \inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - \tau_t|^p \geq (1 + o(1)) \frac{1}{s^p} \inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - X_t^\ast|^p, \tag{16} \]
where the infimum is over estimators that depend only on \( Y_t^\ast \).

Since \( (X_t^\ast, Y_t^\ast) \) are jointly Gaussian, for all \( p \geq 1 \) the infimum on the right-hand side of (16) is achieved by \( \hat{X}_t \), the mmse estimator (3) of \( X_t \) given observation \( Y_t^\ast \). It then follows that
\[ \inf_{\eta(Y_t^\ast)} \mathbb{E}|\eta - X_t^\ast|^p = \left( \frac{\ell \varepsilon^2}{s(1 + \varepsilon^2)} \right)^{p/2} \mathbb{E}|N|^p \]
which, together with (16), gives (14).

To achieve the right-hand side of (14), it is natural to consider the stopping time
\[ \eta^\circ_t = t^* + \frac{(\ell - \hat{X}_t^*)^+}{s} \tag{17} \]
which is similar to the right-hand side expression of (15), except that \( X_t \) is replaced by its (optimal) mmse estimator \( \hat{X}_t \) (the discrepancy due to the rounding in (17) plays no role asymptotically).

This stopping time is in fact optimal since the moments of \( \eta^\circ_t - \eta^\circ_t \) coincide with the right-hand side of (14), asymptotically. Finally, since \( \hat{X}_t \) is the best estimator of \( X_t \), \( \eta^\circ_t \) also represents a natural candidate since it is based on sequentially estimating \( X \) in an optimal fashion.

We proceed with the formal proof. Lower bound: Fix \( p \geq 1 \) and fix an integer \( t \geq 1 \)—later we take \( t = t^* \) defined in (5).

Then,
\[ (\inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - \tau_t|^p)^{1/p} = \left( \inf_{\eta(Y_0^\infty)} \mathbb{E} \left| \left( \eta - t - \frac{\ell - X_t}{s} \right) - \left( \tau_t - t - \frac{\ell - X_t}{s} \right) \right|^p \right)^{1/p} \geq \left( \inf_{\eta(Y_0^\infty)} \mathbb{E} \left| \eta - t - \frac{\ell - X_t}{s} \right|^p \right)^{1/p} - \left( \mathbb{E} \left| \tau_t - t - \frac{\ell - X_t}{s} \right|^p \right)^{1/p} \]
\[ = \left( \inf_{\eta(Y_t^\ast)} \mathbb{E} \left| \eta - t - \frac{\ell - X_t}{s} \right|^p \right)^{1/p} - \left( \mathbb{E} \left| \tau_t - t - \frac{\ell - X_t}{s} \right|^p \right)^{1/p}, \tag{18} \]
where the inequality holds by the triangle inequality, and where the last equality holds since \( Y_t^\ast \) is a sufficient statistics for \( X_t \).

Since \( (X_t, Y_t^\ast) \) are jointly Gaussian,
\[ X_t \overset{d}{=} \hat{X}_t + \left( \frac{\ell \varepsilon^2}{s(1 + \varepsilon^2)} \right)^{1/2} N, \tag{19} \]
where \( \hat{X}_t \) is the mmse estimator of \( X_t \) given observation \( Y_t^\ast \) defined in (3), and where \( N \sim N(0, 1) \) is independent of \( \hat{X}_t \). Hence,
\[ \inf_{\eta(Y_t^\ast)} \mathbb{E} \left| \eta - t - \frac{\ell - X_t}{s} \right|^p = \frac{1}{s^p} \inf_{\eta(Y_t^\ast)} \mathbb{E} |\eta s - t s - \ell - X_t|^p = \frac{1}{s^p} \mathbb{E} |\hat{X}_t - X_t|^p = \frac{1}{s^p} \mathbb{E} \left| \hat{X}_t - X_t \right|^p = \left( \frac{\ell \varepsilon^2}{s^2(1 + \varepsilon^2)} \right)^{p/2} \mathbb{E} |N|^p. \tag{20} \]
any absolute moment with respect to $X_t$. The fourth equality holds by (19).

We now upperbound the second term on the right-hand side of (18). As we shall see, compared to the first term, the contribution of the second term is negligible when $t = t^*$.

We have
\[
\mathbb{E}[(\tau_\ell - t) - (\ell - X_t)/s]\P^p
= \mathbb{E}[(\tau_\ell - t) - (\ell - X_t)/s; \tau_\ell \leq t]
+ \mathbb{E}[(\tau_\ell - t) - (\ell - X_t)/s; \tau_\ell > t].
\tag{21}
\]

For the first term on the right-hand side of (21),
\[
\mathbb{E}[(\tau_\ell - t) - (\ell - X_t)/s; \tau_\ell \leq t]
\leq \mathbb{E}[(t + \ell/s + |X_t|/s; \tau_\ell \leq t) \P(\tau_\ell > t)]
\leq \mathbb{E}(k_1 \P(k_2 + (\ell - X_t)_+^{p/2})^{1/2}
\tag{22}
\]
by the triangle inequality and Cauchy-Schwartz inequality, respectively.

For the second term on the right-hand side of (21),
\[
\mathbb{E}[(\tau_\ell - t) - (\ell - X_t)/s; \tau_\ell > t]
= \mathbb{E}[(\tau_\ell - t) - (\ell - X_t)/s; \tau_\ell > t] \P(\tau_\ell > t)
\leq \mathbb{E}(k_1 \P(k_2 + (\ell - X_t)_+^{p/2})^{1/2}
\tag{23}
\]
where the second inequality follows from Claim ii. of Lemma 1 and the strong Markov property of $X_t$ at time $t$, with $k_1, k_2 \geq 0$ being constants that depend only on $p$ and $s$. Combining (18), (20), (21), and (22) yields
\[
\inf_{\eta(Y_{\infty}^\omega)} \mathbb{E}|\eta - \tau_\ell|^{1/p} \geq E[\tau_\ell^{p/2}] \mathbb{E}[\left| X_t \right|^p]
- \left[ \frac{t \epsilon^2}{s^{3/2}} \right]^{p/2} \mathbb{E}[N]^p
+ k_1 \mathbb{E}[k_2 + (\ell - X_t)_+^{p/2})^{1/2}.
\tag{24}
\]
Finally, letting $t = t^*$ where $t^*$ is defined in [5], we have
\[
\P(\tau_\ell \leq t^*) \leq \exp(-\Omega(t^{2p-1}))
\]
by Claim i. of Lemma 1. Therefore,
\[
\mathbb{E}(t^* + \ell/s + |X_{t^*}|/s)^{p/2} \P(\tau_\ell \leq t^*) = o(1) \quad (\ell \to \infty)
\tag{25}
\]
since
\[
X_{t^*} \overset{d}{=} s \cdot t^* + (t^*)^{1/2} N.
\tag{26}
\]
From (26) and (5) we also get
\[
\mathbb{E}(k_2 + (\ell - X_{t^*})_+^{p/2})^{1/2} = O(\ell^{p/2})
= o(\ell^{p/2})
\tag{27}
\]
since $q < 1$. From (24) with $t = t^*$, (25), and (27) we get
\[
\inf_{\eta(Y_{\infty}^\omega)} \mathbb{E}|\eta - \tau_\ell|^{1/p} \geq (1 + o(1)) \left( \frac{\ell \epsilon^2}{s^3(1 + \epsilon^2)} \right)^{p/2} \mathbb{E}[N]^p
\tag{28}
\]
\footnote{\Omega(\cdot) refers to standard order notations, see, e.g., [3] Chapter 3.}

as $\ell \to \infty$, yielding the desired result.

Next, we establish the asymptotic optimality of $\eta^\omega$ and $\eta^\star$ by showing that their absolute moments with respect to $\tau_\ell$ are equal to the right-hand side of (28). The proof of optimality of $\eta^\omega$ uses most of the arguments of the proofs of [2, Theorem 2.1], which establishes optimality of $\eta^\omega$ for $p = 1$, together with some of the arguments used to establish optimality of $\eta^\star$.

Achievability, $\eta^\star$: To simplify exposition, we ignore discrepancies due to the rounding of non-integer quantities as they play no role asymptotically. In particular, we assume that $\eta^\star$ is given by
\[
\eta^\star_\ell = t^* + (\ell - \hat{X}_{t^*})_+/s
\]
without rounding the fraction. Notice that if $\eta^\star$, as defined above, is asymptotically optimal, then a triangle inequality argument immediately shows that $\eta^\star$ with the rounding of the fraction is also asymptotically optimal.

Let
\[
\Delta \overset{\text{def}}{=} \tau_\ell - t^*,
\tag{29}
\]
and let
\[
\hat{\Delta} \overset{\text{def}}{=} (\ell - \hat{X}_{t^*})_+/s.
\tag{30}
\]
Then,
\[
\mathbb{E}|\eta^\star_\ell - \tau_\ell|^{p} = \mathbb{E}[(\hat{\Delta} - \Delta)^{p}; \tau_\ell > t^*]
+ \mathbb{E}[(\eta^\star - \tau_\ell)^{p}; \tau_\ell \leq t^*].
\tag{31}
\]
For the first term on the right-hand side of (31),
\[
\mathbb{E}[(\hat{\Delta} - \Delta)^{p}; \tau_\ell > t^*] \leq \mathbb{E}[(\hat{\Delta} - \Delta)^{p}; \hat{X}_{t^*} < \ell, \tau_\ell > t^*]
+ \mathbb{E}[(\Delta)^{p}; \hat{X}_{t^*} \geq \ell].
\tag{32}
\]
By the triangle inequality,
\[
(\mathbb{E}[(\hat{\Delta} - \Delta)^{p}; \hat{X}_{t^*} < \ell, \tau_\ell > t^*])^{1/p}
\leq (\mathbb{E}[(\hat{\Delta} - (\ell - X_{t^*})/s)^{p}; \hat{X}_{t^*} < \ell, \tau_\ell > t^*])^{1/p}
+ (\mathbb{E}[(\ell - X_{t^*})/s - \Delta)^{p}; \hat{X}_{t^*} < \ell, \tau_\ell > t^*])^{1/p}
\leq (\mathbb{E}[(\hat{\Delta} - (\ell - X_{t^*})/s)^{p}; \hat{X}_{t^*} < \ell])^{1/p}
+ (\mathbb{E}[(\ell - X_{t^*})/s - \Delta)^{p}; \tau_\ell > t^*])^{1/p}.
\tag{33}
\]
For the first term on the right-hand side of (33),
\[
\mathbb{E}[(\hat{\Delta} - (\ell - X_{t^*})/s)^{p}; \hat{X}_{t^*} < \ell]
= \mathbb{E}[(X_{t^*} - \hat{X}_{t^*})/s; \hat{X}_{t^*} < \ell]
\leq |\mathbb{E}[(X_{t^*} - \hat{X}_{t^*})/s]^p|
= \frac{t^* \epsilon^2}{s^{3}(1 + \epsilon^2)^{p/2}} \mathbb{E}[N]^p
\tag{34}
\]
where the last equality follows from (19).

For the second term on the right-hand side of (33) we use (23) with $t = t^*$ to get
\[
\mathbb{E}[(\ell - X_{t^*})/s - \Delta)^{p}; \tau_\ell > t^*] \leq k_1 \mathbb{E}[k_2 + (\ell - X_{t^*})_+^{p/2}/s
\tag{35}
\]
where $k_1, k_2$ are constants that depend on $p$ and $s$ only. 

\footnote{As such, $\eta^\star$ is no more a stopping time, strictly speaking.}
For the second term on the right-hand side of (32), Cauchy-Schwarz inequality yields
\[ \mathbb{E}(|\Delta|^{p}; \hat{X}_{t*} \geq \ell) \leq \left( \mathbb{E}(|\Delta|^{2p}) \right)^{1/2} \mathbb{P}(\hat{X}_{t*} \geq \ell)^{1/2}. \] (36)

By the triangle inequality,
\[ \left( \mathbb{E}(|\Delta|^{2})^{1/2} \right)^{p} \leq \left( \mathbb{E}(|\tau - \ell|/s)^{2p} + \mathbb{E}(|\ell - s - t^{*}|^{2})^{1/2} \right)^{1/2}, \]
\[ \leq k_{1}(k_{2} + \ell)^{1/2} + (\ell/s) q, \] (37)

where for the second inequality we used Claim ii. of Lemma 1 with \( k_{1}, k_{2} \) constants that depend on \( p \) and \( s \), and the definition of \( t^{*} \) (recall that we ignore discrepancies due to the rounding of non-integer quantities).

From (32), (33), (34), (35), (36), and (37) we obtain
\[ \mathbb{E}(|\Delta|^{p}; \tau_{\ell} > t^{*}) \leq \left[ \left( \frac{1}{s^{p}} \frac{t^{*}}{1 + \varepsilon^{2}} \right)^{p/2} \mathbb{E}[|N|^{p}] \right]^{1/2}, \]
\[ + \left[ k_{1} \mathbb{E}\left( k_{2} + (\ell - X_{t*})^{+} \right)^{p/2} \right]^{1/2}, \]
\[ + \left[ k_{1}(k_{2} + \ell)^{1/2} + (\ell/s)^{q} \right]^{p} \mathbb{P}(\hat{X}_{t*} \geq \ell)^{1/2}. \] (38)

For the second term on the right-hand side of (31), using Cauchy-Schwarz inequality and the triangle inequality we get
\[ \mathbb{E}(|\tau_{\ell} - \eta_{\ell}^{*}|^{p}; \tau_{\ell} \leq t^{*}) \leq \left[ \left( \mathbb{E}(|\tau_{\ell} - \ell|/s)^{2p} + \mathbb{E}(\ell - s - \eta_{\ell}^{*})^{2p} \right)^{1/2} \right]^{1/2}, \]
\[ \leq \left[ k_{1}(k_{2} + \ell)^{1/2} + k_{3}(k_{4} + \ell)^{1/2} \right]^{p} \mathbb{P}(\tau_{\ell} \leq t^{*})^{1/2}, \] (39)

where for the second inequality we used Claim ii. of Lemma 1 with \( k_{3}, k_{4} \) constants that depend on \( p, s, \) and \( \varepsilon \).

Combining (31), (38), and (39) we obtain
\[ \mathbb{E}[|\eta_{\ell}^{*} - \tau_{\ell}|^{p}] \leq \left[ \left( \frac{1}{s^{p}} \frac{t^{*}}{1 + \varepsilon^{2}} \right)^{p/2} \mathbb{E}[|N|^{p}] \right]^{1/2}, \]
\[ + \left[ k_{1} \mathbb{E}\left( k_{2} + (\ell - X_{t*})^{+} \right)^{p/2} \right]^{1/2}, \]
\[ + \left[ k_{1}(k_{2} + \ell)^{1/2} + (\ell/s)^{q} \right]^{p} \mathbb{P}(\hat{X}_{t*} \geq \ell)^{1/2}, \]
\[ + \left[ k_{1}(k_{2} + \ell)^{1/2} + k_{3}(k_{4} + \ell)^{1/2} \right]^{p} \mathbb{P}(\tau_{\ell} \leq t^{*})^{1/2}. \] (40)

Using (3) and Claim i. of Lemma 1 one deduces that the third and fourth terms on the right-hand side of (40) tend to zero as \( \ell \to \infty \). Since \( X_{t*} \approx s \cdot t^{*} + (t^{*})^{1/2} N \) and \( t^{*} = (\ell/s)(1 + o(1)) \), we conclude that
\[ \mathbb{E}[|\eta_{\ell}^{*} - \tau_{\ell}|^{p}] \leq \left( 1 + o(1) \right) \left( \frac{\ell}{s^{3}} \mathbb{E}[|N|^{p}] \right)^{1/2}, \]
\[ = \left( 1 + o(1) \right) C_{1}(\ell, s, \varepsilon, p). \]

As \( \ell \to \infty \), where
\[ C_{1}(\ell, s, \varepsilon, p) \overset{\text{def}}{=} \left( \frac{\ell}{s^{3}} \right)^{p/2} \mathbb{E}[|N|^{p}]. \]

This establishes the asymptotic optimality of \( \eta_{\ell}^{*} \).

**Achievability, \( \eta_{\ell}^{\circ} \):** We write \( \mathbb{E}[|\eta_{\ell}^{\circ} - \tau_{\ell}|^{p}] \) as
\[ \mathbb{E}[|\eta_{\ell}^{\circ} - \tau_{\ell}|^{p}] = \mathbb{E}[|\eta_{\ell}^{\circ} - \tau_{\ell}|^{p}; \eta_{\ell}^{\circ} \geq \tau_{\ell} \}, \]
\[ + \mathbb{E}[|\tau_{\ell} - \eta_{\ell}^{\circ}|^{p}; \tau_{\ell} \geq \eta_{\ell}^{\circ} \}, \] (41)

and upper bound each of the two terms on right-hand side of the above equation. As in the previous section, we ignore discrepancies due to the rounding of non-integer quantities as they play no role asymptotically. In particular, we treat \( \ell/s \) as an integer.

Letting
\[ \nu \overset{\text{def}}{=} \inf \{ t \geq 0 : \hat{X}_{\tau-t} \geq \ell \}, \]

we have
\[ \mathbb{E}[|\eta_{\ell}^{\circ} - \tau_{\ell}|^{p}; \eta_{\ell}^{\circ} \geq \tau_{\ell} \}
\[ = \mathbb{E}(\nu^{p}; \eta_{\ell}^{\circ} \geq \tau_{\ell} \}
\[ \leq \left[ \mathbb{E}(\ell - \hat{X}_{\tau_{\ell}}; \hat{X}_{\tau_{\ell}} < \ell)^{1/p} \right]^{p}, \]
\[ + \left[ \mathbb{E}(\nu^{p} - (\ell - \hat{X}_{\tau_{\ell}})/s; \hat{X}_{\tau_{\ell}} < \ell)^{1/p} \right]^{p}, \] (42)

where the first inequality follows from the definition of \( \hat{X}_{\ell} \) (see (2)) and where the second inequality follows from the triangle inequality.

We upper bound the two expectations on the right-hand side of (42).

For the first term, for \( i \geq 1 \) let
\[ U_{i} \overset{\text{def}}{=} (X_{i} - \hat{X}_{i}) - (X_{i-1} - \hat{X}_{i-1}) \]
\[ = \left( \varepsilon^{2}/(1 + \varepsilon^{2}) \right) V_{i} - \varepsilon \left( (1 + \varepsilon^{2}) \right) W_{i} \] (43)
\[ \overset{\text{def}}{=} \varepsilon \left( 1 + \varepsilon^{2} \right)^{1/2} N. \] (44)

Then
\[ X_{\tau_{\ell}} - \hat{X}_{\tau_{\ell}} = \sum_{i=1}^{\ell/s} U_{i} - \mathbb{I} \{ \tau_{\ell} < \ell/s \} \sum_{i=\tau_{\ell}+1}^{\ell/s} U_{i} \]
\[ + \mathbb{I} \{ \tau_{\ell} > m \} \sum_{i=\ell/s+1}^{\ell/s} U_{i}, \] (45)

where \( \mathbb{I} \{ \cdot \} \) denotes the indicator function of event \( \mathcal{A} \).
and, by the triangle inequality\textsuperscript{19}

\[
\left[ \mathbb{E}(X_{\tau_\ell} - \hat{X}_{\tau_\ell})^p \right]^{1/p} \leq \left[ \mathbb{E} \left( \left( \sum_{i=1}^{\ell/s} U_i \right)^p \right) \right]^{1/p} + \left[ \mathbb{E} \left( - \mathbb{I} \{\tau_\ell < \ell/s\} \sum_{i=\tau_\ell+1}^{\ell/s} U_i \right)^p \right]^{1/p} + \left[ \mathbb{E} \left( \mathbb{I} \{\tau_\ell > \ell/s\} \sum_{i=\ell/s+1}^{\tau_\ell} U_i \right)^p \right]^{1/p}.
\]

(46)

We bound each term on the right-side of (46). For the first term, from (44) we have

\[
\mathbb{E} \left( \sum_{i=1}^{\ell/s} U_i \right)^p = \left( (\ell/s)\varepsilon^2/(1+\varepsilon^2) \right)^{p/2} \mathbb{E} N_+^p.
\]

(47)

For the second term on the right-side of (46), using (44), together with the fact that \(\tau_\ell\) is independent of \(U_{\tau_\ell+1}, U_{\tau_\ell+2}, \ldots\) we get

\[
\mathbb{E} \left( - \mathbb{I} \{\tau_\ell < \ell/s\} \sum_{i=\tau_\ell+1}^{\ell/s} U_i \right)^p = \mathbb{E} \left[ (\ell/s-\tau_\ell) + \varepsilon^2/(1+\varepsilon^2) \right]^{p/2} \mathbb{E} N_+^p \leq (\mathbb{E} |\ell/s-\tau_\ell|^{p/2}) \mathbb{E} N_+^p \leq k_1 (k_2 + \varepsilon)^{p/4} \mathbb{E} N_+^p = O(\ell^{p/4}),
\]

(48)

where for the first inequality we bounded \(\varepsilon^2/(1+\varepsilon^2)\) by 1, and where for the second inequality we used Claim ii. of Lemma\textsuperscript{11}. For the third term on the right-side of (46), using (43), the triangle inequality, and by upperbounding \(\varepsilon^2/(1+\varepsilon^2)\) and \(\varepsilon/(1+\varepsilon^2)\) by 1, we get

\[
\left[ \mathbb{E} \left( \mathbb{I} \{\tau_\ell > \ell/s\} \sum_{i=\ell/s+1}^{\tau_\ell} U_i \right)^p \right]^{1/p} \leq \mathbb{E} \left( \left( \sum_{i=\ell/s+1}^{\tau_\ell} W_i \right)^p \right) \left( \sum_{i=\ell/s+1}^{\tau_\ell} W_i \right) \mathbb{E} N_+^p + \left( \sum_{i=\ell/s+1}^{\tau_\ell} V_i \right)^p \mathbb{E} N_+^p.
\]

(49)

Since \(\tau_\ell\) and \(\{W_i\}\) are independent, we have

\[
\sum_{i=\ell/s+1}^{\tau_\ell} W_i \overset{d}{=} \sqrt{(\tau_\ell - \ell/s)} + N,
\]

and a similar calculation as for (48) shows that

\[
\mathbb{E} \left( \sum_{i=\ell/s+1}^{\tau_\ell} W_i \right)^p = O(\ell^{p/4}).
\]

(50)

\textsuperscript{19}By \(x_i^p\) we actually mean \((x_+)^p\).

We now focus on the second expectation on the right-side of (49). Since, on \(\{\tau_\ell > \ell/s\}\), we have

\[
\sum_{i=\ell/s+1}^{\tau_\ell} V_i = (X_{\tau_\ell} - X_{\ell/s}) - s(\tau_\ell - \ell/s),
\]

we consider the shifted process \(\{X_\ell - X_{\ell/s}\}_{\ell \geq \ell/s}\) and its crossing of level \(\ell - X_{\ell/s}\). It then follows that

\[
\mathbb{E} \left( \mathbb{I} \{\tau_\ell > \ell/s\} \sum_{i=\ell/s+1}^{\tau_\ell} V_i \right)^2
\]

\[
= s^2 \mathbb{E} \left( \left( (X_{\tau_\ell} - X_{\ell/s})/s - (\tau_\ell - \ell/s) \right)^2 \right) \leq s^2 \mathbb{E} \left[ \left( (X_{\tau_\ell} - X_{\ell/s})/s - (\tau_\ell - \ell/s) \right)^2 \right] \leq k_1 \mathbb{E} (k_2 + (X_{\tau_\ell} - X_{\ell/s})/s)^{p/2} = O(\ell^{p/4}),
\]

(51)

where \(k_1, k_2\) are constants that depend only on \(s\) and \(p\), and where the second inequality follows Claim ii. of Lemma\textsuperscript{11} and the Markov property of process \(X\) at time \(\ell/s\). We now justify the second equality in (51). We have

\[
X_{\ell/s} \overset{d}{=} \ell + \sqrt{sN}
\]

and

\[
X_{\tau_\ell} = \ell + \varepsilon_{\tau_\ell},
\]

where \(\varepsilon_{\tau_\ell}\) denotes the excess over the boundary at \(\tau_\ell\). Using this and the triangle inequality we get

\[
\left( \mathbb{E} \left( (X_{\tau_\ell} - X_{\ell/s})^2 \right)^{p/2} \right)^{2/p} = \left( \mathbb{E} \varepsilon_{\tau_\ell}^{p/2} \right)^{2/p} + \sqrt{s \mathbb{E} N_+^{p/2}}^{2/p},
\]

(52)

which implies that

\[
\mathbb{E} \left( (X_{\tau_\ell} - X_{\ell/s})^2 \right)^{p/2} = O(\ell^{p/4})
\]

since \(\mathbb{E} \varepsilon_{\tau_\ell}^{p/2}\) can be upper bounded by a finite constant that is independent of \(\ell\) (Equation (2))). This establishes the second equality in (51). Combining (49) together with (50) and (51) yields

\[
\mathbb{E} \left( \mathbb{I} \{\tau > \ell/s\} \sum_{i=\ell/s+1}^{\tau} U_i \right)^p = O(\ell^{p/4}).
\]

(53)

From (46), (47), (48), and (53) we get

\[
\mathbb{E} \left( X_{\tau_\ell} - \hat{X}_{\tau_\ell} \right)^p \leq (1 + o(1)) \left( \frac{\ell^2}{(1+\varepsilon^2)} \right)^{p/2} \mathbb{E} N_+^p.
\]

(54)

For the second expectation on the right-hand side of (42) we have

\[
\mathbb{E} [\nu - \hat{X}_{\tau_\ell}]/s \mathbb{E} \left( \hat{X}_{\tau_\ell} - \ell \right] \leq k_3 \mathbb{E} [k_4 + \ell - \hat{X}_{\tau_\ell}]^{p/4} = O(\ell^{p/4}),
\]

(55)

where the inequality follows from the strong Markov property of \(\hat{X}\) at time \(\tau_\ell\) together with Claim ii. of Lemma\textsuperscript{11} with \(k_3\) and \(k_4\) constants that depend on \(s\) and \(\varepsilon\). From (42), (54), and (55) we get

\[
\mathbb{E} (\eta^0_\ell - \tau_\ell; \eta^0_\ell \geq \tau_\ell) \leq (1 + o(1)) \left( \frac{\ell^2}{s^3(1+\varepsilon^2)} \right)^{p/2} \mathbb{E} N_+^p.
\]

(56)
Using analogous arguments as for establishing (56), which essentially amounts to swap the roles of $X$ and $X$ and the roles of $\tau$ and $\eta^*_d$, we get

$$\mathbb{E}(\tau_\ell - \eta^*_{d} \mid p; \tau \geq \eta^*_{d}) \leq (1 + o(1)) \left( \frac{\ell \varepsilon^2}{s^3(1 + \varepsilon^2)} \right)^{p/2} \mathbb{E}N^{p}_{\ell}.$$  

(57)

Finally, from (41), (56), and (57) we get

$$\mathbb{E}|\tau_\ell - \eta^*_{d} |^p \leq (1 + o(1)) \left( \frac{\ell \varepsilon^2}{s^3(1 + \varepsilon^2)} \right)^{p/2} \mathbb{E}|N|^{p}_{\ell} \quad (\ell \to \infty),$$

which establishes the asymptotic optimality of $\eta^*_{d}$.

**B. Proof of Theorem 4**

As mentioned earlier, $\eta^*_{d}$ is a very natural stopping time to consider since, on average, $X_\ell$ is $s \cdot d$ higher than $Y_t$. Now, the time needed to go from level $\ell - s \cdot d$ to level $\ell$ has (approximately) the Gaussian distribution $d + (\sqrt{d} / s) N$ by Claim iii. of Lemma 1. Hence we have $\tau_\ell - \eta^*_{d} \approx (\sqrt{d} / s) N$ which yields the second equality in Theorem 4. The optimality of $\eta^*_{d}$ is established essentially by showing that any (asymptotically) optimal stopping rule shouldn’t stop later than $\eta^*_{d}$.

**Lower bound:** Let $\ell$ be any function of $d$ such that $\ell \geq s \cdot d$, and fix integer $d \geq 1$. Further, let

$$\nu = \inf \{ t \geq 0 : X_t \geq \ell - s \cdot d (1 - \varepsilon) \}$$

where $\varepsilon$ is a constant such that $0 < \varepsilon < 1$—later we take $\varepsilon \to 0$.

Then,

$$\inf_{\eta} \mathbb{E} |\eta - \tau_\ell |^p \geq \mathbb{E} \left( |\eta - \tau_\ell |^p; \tau_\ell \leq \nu + d \right) \geq \inf_{\eta(Y^{\nu+d} \leq \nu+d)} \mathbb{E} \left( |\eta - \tau_\ell |^p; \tau_\ell \leq \nu + d \right)$$

$$\geq \inf_{\eta(X^{\nu+d} \leq \nu+d)} \mathbb{E} \left( |\eta - \tau_\ell |^p; \tau_\ell \geq \nu + d \right),$$

(58)

where the infimum on the right-hand side of the second inequality is over all estimators that depend on $Y^{\nu+d}$ (these estimators need not be stopping times), and where the equality holds since $Y_t = X_{t-d}$.

Let

$$\delta = \inf \{ t \geq 0 : X_{t+d} \geq \ell \},$$

so that, by definition,

$$\tau_\ell = \nu + \delta.$$

Then,

$$\inf_{\eta(X^{\nu+d} \leq \nu+d)} \mathbb{E} \left( |\eta - \tau_\ell |^p; \tau_\ell \leq \nu + d \right)$$

$$= \inf_{\eta(Y^{\nu+d} \leq \nu+d)} \mathbb{E} \left( |\eta - (\nu + \delta) |^p; 0 \leq \delta \leq \nu + d \right)$$

$$= \inf_{\eta(Y^{\nu+d} \leq \nu+d)} \mathbb{E} \left( |\eta - \delta |^p; 0 \leq \delta \leq d \right)$$

$$= \inf_{\eta(X^{\nu+d} \leq \nu+d)} \mathbb{E} \left( |\eta - \delta |^p; 0 \leq \delta \leq d \right)$$

$$\geq \inf_{\eta(X^\nu) \leq \nu+d} \mathbb{E} \left( |\eta - \delta |^p; 0 \leq \delta \leq d \right)$$

$$\geq \inf_{\eta(X^\nu)} \mathbb{E} \left( |\eta - \delta |^p; 0 \leq \delta \leq d, \varepsilon \leq c \right).$$

(59)

The second equality in (59) follows from Fact 1. The infimum on the right-hand side of the third equality is over estimators that depend on $X_\nu$ only, since $\delta$ is defined over $X_{\nu}, X_{\nu+1}, \ldots$,.

The last inequality holds for an arbitrary fixed constant $c > 0$, with $\varepsilon$ defined as the excess at time $\nu$, i.e.,

$$\varepsilon = \inf \left( X_\nu - (\ell - s \cdot d (1 - \varepsilon)) \right) \geq 0.$$

Take $d$ large enough so that

$$\frac{sd \varepsilon}{d \varepsilon} > c,$$

and define

$$d \varepsilon = d(1 - \varepsilon) + \varepsilon,$$

$$N_\nu = \sqrt{\frac{d}{d \varepsilon}} \left( \delta - d \varepsilon \right),$$

and define the functions $f_1(d, \varepsilon)$ and $f_2(d, \varepsilon)$ as

$$f_1(d, \varepsilon) = \frac{s \sqrt{d(1 - \varepsilon)}}{d \varepsilon}$$

and

$$f_2(d, \varepsilon) = \frac{sd \varepsilon - c}{d(1 - \varepsilon) + c/s}.$$

Notice that both $f_1$ and $f_2$ are strictly positive because of (60). Using the definitions of $d \varepsilon$ and $N_\nu$ we get

$$\inf_{\eta(X_\nu)} \mathbb{E} \left( |\eta - \delta |^p; 0 \leq \delta \leq d, \varepsilon \leq c \right)$$

$$= \inf_{\eta(X_\nu)} \mathbb{E} \left( |\eta - (\delta - d \varepsilon) |^p; \varepsilon \leq c \right)$$

$$\geq \inf_{\eta(X_\nu)} \mathbb{E} \left( |\eta - (\delta - d \varepsilon) |^p; \varepsilon \leq c \right)$$

$$= \inf_{\eta(X_\nu)} \mathbb{E} \left( |\eta - \delta |^p; 0 \leq \delta \leq d, \varepsilon \leq c \right).$$

(61)

where we defined the events

$$\varepsilon_1 = \{ -s \sqrt{d \varepsilon} \leq N_\nu \leq s (d \varepsilon) / \sqrt{d \varepsilon}, \varepsilon \leq c \}$$

$$\varepsilon_2 = \{ -f_1(d, \varepsilon) \leq N_\nu \leq f_2(d, \varepsilon), \varepsilon \leq c \}.$$

The first equality in (61) holds by Fact 1. The first inequality holds by the definitions of $f_1(d, \varepsilon)$ and $f_2(d, \varepsilon)$ and by noting that, on $\{ \varepsilon \leq c \}$, the range of $N_\nu$ in $\varepsilon_2$ contains the range of $N_\nu$ in $\varepsilon_2$. The second inequality holds by the definition of $N_\nu$ and because on event $\varepsilon_2$ we have

$$d \varepsilon \geq d(1 - \varepsilon).$$

Finally the last equality in (61) holds by Fact 1 since $d \varepsilon$ is a function of $X_\nu$ (through $\varepsilon$).

Since $f_1(d, \varepsilon)$ and $f_2(d, \varepsilon)$ are increasing functions of $d$, let us pick $d$ so that the following inequality, more stringent than (60), is satisfied

$$c < \min \{ s \varepsilon; f_1(d, \varepsilon), f_2(d, \varepsilon) \}.$$

(62)

It then follows that

$$\mathbb{E} \left( |\eta - N_\nu |^p; -f_1(d, \varepsilon) \leq N_\nu \leq f_2(d, \varepsilon), \varepsilon \leq c \right)$$

$$\geq \mathbb{E} \left( |\eta - N_\nu |^p; -c \leq N_\nu \leq c, \varepsilon \leq c \right).$$
hence, from (61),
\[
\frac{1}{(d(1-\varepsilon))^2} \inf_{\eta(X_t)} E(|\eta - \delta|^p; 0 \leq \delta \leq d, \epsilon_\nu \leq c) \\
\geq \inf_{\eta(X_t)} E(|\eta - N_\nu|^p; -c \leq N_\nu \leq c, \epsilon_\nu \leq c) \\
= \left[ \inf_{\eta(X_t)} E\left(|\eta - N_\nu|^p; -c \leq N_\nu \leq c| \epsilon_\nu \leq c \right) \right] P(\epsilon_\nu \leq c). 
\]
(63)

Now, \( E\epsilon_\nu \), can be upperbounded by a constant \( 0 \leq k < \infty \) that is independent of the barrier level at time \( \nu \), i.e., \( \ell = s \cdot d(1-\varepsilon) \) (see [7] Equation (2)). Hence,
\[
P(\epsilon_\nu \leq c) \geq 1 - k/c 
\]
by Markov inequality. Therefore, for any fixed \( 0 < \varepsilon < 1 \), \( c \)
large enough so that
\[
k/c \leq \varepsilon \quad (64)
\]
and \( d \) large enough so that (62) holds, from (63) we have
\[
\frac{1}{(1-\varepsilon)} \frac{1}{(d(1-\varepsilon))^2} \inf_{\eta} E|\eta - \tau_\xi|^p \\
\geq \inf_{\eta(X_t)} E\left(|\eta - N_\nu|^p; -c \leq N_\nu \leq c| \epsilon_\nu \leq c \right) \\
\geq E\left(|\eta|^p; -c \leq N \leq c \right) \quad (65)
\]
where the infimum on the right-hand side of the second inequality is over constant estimators, and where the last inequality follows from the symmetry and monotonicity of the probability density function of \( N \) around zero.

Since the above inequality holds for arbitrary \( 0 < \varepsilon < 1 \) and \( c > 0 \) such that (64) is satisfied, by letting \( c = c(\varepsilon) = k/\varepsilon \) and by taking \( \varepsilon \rightarrow 0 \) on both sides of (65) yields
\[
\lim_{d \rightarrow \infty} \frac{1}{(1-\varepsilon)} \frac{1}{(d(1-\varepsilon))^2} \inf_{\eta} E|\eta - \tau_\xi|^p \\
\geq k/c \inf_{\eta(X_t)} E|\eta|^p, 
\]
\[
\inf_{\eta} E|\eta - \tau_\xi|^p \geq E|N|^p \quad (66)
\]

and
\[
\tau_\xi = \xi + \Delta. 
\]

Further, define
\[
\Delta_\omega \triangleq \inf\{t \geq 0 : X_{t+\xi} - X_\eta \geq s d\}. 
\]

Notice that if there were no barrier overshoot at time \( \xi \), then \( X_\xi = \ell - s \cdot d \) and so \( \Delta_\omega \) would be equal to \( \Delta \).

It follows that
\[
E|\eta_\omega^* - \tau_\xi|^p = E|\Delta - d|^p \\
\leq \left[ (E|\Delta_\omega - d|^p)^{1/p} + (E|\Delta_\omega - \Delta|^p)^{1/p} \right]^p \\
= \left[ (E|\Delta_\omega - d|^p)^{1/p} + (E|\Delta_\omega - \Delta|^p)^{1/p} \right]^p 
\]
where
\[
\epsilon_\xi \triangleq X_\xi - (\ell - s \cdot d) 
\]
denotes the excess at time \( \xi \). The first inequality in (66) follows from the triangle inequality and the second inequality follows from the strong Markov property of \( X \) at time \( \xi \).

From Claim iii. of Lemma 1 and the strong Markov property of \( X \) at time \( \xi \),
\[
E|\Delta_\omega - d|^p = (1 + o(1)) \frac{dp/2}{s^p} E|N|^p 
\]
as \( d \rightarrow \infty \).

Assume that \( E\tau_\xi^p \) can be upper bounded by a finite constant that does not depend on \( d \). Then, from (66) and (67) we get
\[
E|\eta_\omega^* - \tau_\xi|^p \leq (1 + o(1)) \frac{dp/2}{s^p} E|N|^p 
\]
as \( d \rightarrow \infty \) while \( \ell \geq s \cdot d \), yielding the desired result.

As we now show, the fact that \( E\tau_\xi^p \) can be upper bounded by a finite constant that does not depend on the barrier level at time \( \eta \), for notational convenience, we drop the subscript \( \xi \) and write \( \epsilon \) in place of \( \epsilon_\xi \).

If the barrier level at time \( \eta \), i.e., \( \ell - s \cdot d \), is bounded in the limit \( d \rightarrow \infty \), i.e., if \( \limsup_{d \rightarrow \infty}(\ell - s \cdot d) < \infty \), then clearly \( E\tau_\xi^p \) can be upper bounded by a finite constant that does not depend on \( d \).

Now, suppose that \( \lim_{d \rightarrow \infty}(\ell - s \cdot d) = \infty \), and suppose, by contradiction, that \( E\tau_\xi^p \rightarrow \infty \). We start with \( p = 1 \).

By Claim ii. of Lemma 1 we have
\[
\tau_\xi = \frac{\epsilon}{s} + \left( \frac{\epsilon}{s} \right)^{1/2} \tilde{N}_\omega \quad (68)
\]
where \( \tilde{N}_\omega \rightarrow N \) in distribution, uniformly over \( \{ \epsilon \geq k \} \), as \( k \rightarrow \infty \). Using this,
\[
E\tau_\xi \leq E(\tau_\xi; \epsilon \leq k) + E(\tau_\xi; \epsilon \geq k) \\
\leq E(\tau_\xi) + E(\epsilon/s + |\epsilon|s^{1/2}) \tilde{N}_\omega; \epsilon \geq k) \\
\leq E(\tau_\xi) + (1/s) E\epsilon + E(\tilde{N}_\omega|\epsilon|s^{1/2}) \\
\leq E(\tau_\xi) + (1/s) E\epsilon + s^{-3/2}(E\epsilon^2)(E(N)^{1/2}) \\
\leq E(\tau_\xi) + (1/s) E\epsilon + s^{-3/2}(E\epsilon^2)(2E(N)^{1/2}) \quad (69)
\]
The first inequality holds since \( \tau_\ell \geq \tau_{\ell'} \) for \( \ell \geq \ell' \). The second inequality follows from (68). The fourth inequality holds by Cauchy-Schwartz inequality. The last inequality holds by (68) for \( k \) large enough.

From (69), if \( E_{\tau_k} \to \infty \) then \( E_e \to \infty \), a contradiction since [7] Equation (2) says that \( E_e \) admits a finite upper bound that does not depend on the barrier level. Hence, \( E_{\tau_k} \to \infty \) can be upper bounded by a finite constant that does not depend on \( d \).

For \( p > 2 \), a similar argument as above shows that \( E_{\tau_k} \to \infty \). In particular, a similar computation as in (68) holds, with the addition of a triangle inequality for the second inequality in (68) to get

\[
E(\tau_k^p; e \geq k) \leq \left( (1/s)(Ee^p)^{1/p} + (E(\hat{N}_k^p[e/s^3]^{p/2}))^{1/p} \right)^p.
\]

This shows for any \( \ell = \ell(d) \geq s \cdot d \), \( \limsup_{d \to \infty} E_{\tau_k}^p < \infty \), yielding the desired result.

C. Proof of Theorem 3

Fix \( p \geq 1/2 \). Suppose for the moment that a stopping time \( \eta \) on \( Y \) that satisfies \( \mathbb{P}(\eta < \tau_{\ell} + d) > 0 \) also satisfies

\[
\mathbb{E}(|\eta - \tau_{\ell}|^p|Y, \eta < \tau_{\ell} + d) = \infty.
\]

Hence, if \( \eta \) satisfies \( \mathbb{E}[|\eta - \tau_{\ell}|^p] < \infty \), then necessarily

\[
\mathbb{P}(\eta \geq \tau_{\ell} + d) = 1.
\]

From this equivalence if follows that

\[
\inf_{\eta} \mathbb{E}[|\eta - \tau_{\ell}|^p] = \inf_{\eta; p(\eta \geq \tau_{\ell} + d) = 1} \mathbb{E}[|\eta - \tau_{\ell}|^p] \geq d^p = \mathbb{E}[\eta_d^*-\tau_{\ell}]^p,
\]

where \( \eta_d^* = \inf\{t \geq 0 : Y_t \geq \ell\} \). Therefore we have the desired result

\[
\inf_{\eta} \mathbb{E}[|\eta - \tau_{\ell}|^p] = d^p = \mathbb{E}[\eta_d^*-\tau_{\ell}]^p.
\]

We prove (70) assuming \( \mathbb{P}(\eta < \tau_{\ell} + d) > 0 \). Equivalently, we show that for any stopping rule \( \eta \) over \( X \) (instead of \( Y \)) such that \( \mathbb{P}(\eta < \tau_{\ell}) > 0 \), necessarily we have

\[
\mathbb{E}[|\eta - \tau_{\ell}|^p|X, \eta < \tau_{\ell}] = \infty.
\]

Given \( X_\eta = \ell - h \), for some arbitrarily fixed \( h > 0 \), let \( \{B_t\}_{t \geq 0} \) be the continuous time version of \( X \) starting at time \( \eta \), i.e., \( \{B_t\}_{t \geq 0} \) is a standard Wiener process starting at time \( \eta \) at level \( B_0 = \ell - h \) and such that \( B_t = X_{\eta+t} \) for \( t = 0, 1, 2, \ldots \).

Let

\[
\tau_{\ell} \explain{def}{=} \inf\{t \geq 0 : B_t = \ell\}.
\]

Suppose \( \eta < \tau_{\ell} \). Since \( \tau_{\ell} \leq \tau_{\ell} - \eta \), had we proved that \( E_{\tau_{\ell}}^p = \infty \), (71) would hold.

From the reflection principle

\[
\mathbb{P}(\tau_{\ell} \leq t) = 2\mathbb{P}(B_t \geq h) = 2Q \left( \frac{h}{\sqrt{t}} \right) \quad h > 0, \quad t > 0,
\]

where \( Q(x) = (1/\sqrt{2\pi}) \int_x^\infty \exp(-x^2/2)dx \). Hence,

\[
E_{\tau_{\ell}}^p = 2 \int_0^\infty tpQ \left( \frac{h}{\sqrt{t}} \right) \ dt = \frac{h}{\sqrt{2\pi}} \int_0^\infty \frac{t^p}{t^{3/2}} e^{-h^2/2t} dt
\]

\[
> \frac{he^{-h^2/2}}{\sqrt{2\pi}} \int_0^\infty \frac{t^p}{t^{3/2}} dt.
\]

Therefore, if \( p \geq 1/2 \), then \( E_{\tau_{\ell}}^p = \infty \), yielding the desired result.

D. Proof of Lemma 7

Claim i. For any real constant \( q \), \( S_t = \sum_{i=1}^t Z_i \) satisfies

\[
\mathbb{E}[\exp^{qS_t+qs\sigma^2/2}] \leq e^{qS_t+qs\sigma^2/2}
\]

which can readily be checked by direct computation.

Hence, letting

\[
M_t = e^{qS_t-rt} \quad t \geq 1
\]

where \( r \) is an arbitrary constant, we get

\[
\mathbb{E}[M_{t+1}|M_1, \ldots, M_t] = M_t e^{qS_t+qs\sigma^2/2-r} \quad t \geq 1.
\]

Let us set \( r = qs + q^2\sigma^2 \) so that

\[
M_t = e^{qS_t-(qs+q^2\sigma^2/2)t} \quad t \geq 1
\]

is a martingale, and introduce the stopping time

\[
\tau_{\ell} = \min\{\lfloor k \rfloor, \tau_{\ell}\}
\]

where \( k > 0 \) is an arbitrary constant. It follows that

\[
1 = \mathbb{E}M_1 = \mathbb{E}M_{\tau_{\ell}} \\
\geq \mathbb{E}[M_{\tau_{\ell}}|\tau_{\ell} < k] \\
\geq e^{-q\ell - (qs+q^2\sigma^2/2)k} \mathbb{P}(\tau_{\ell} < k) \quad q \geq 0 \\
= e^{-q\ell - (qs+q^2\sigma^2/2)k} \mathbb{P}(\tau_{\ell} < k),
\]

where the second equality follows from Doob’s stopping theorem and where the second inequality is valid for \( q \geq 0 \) since \( S_{\tau_{\ell}} \geq \ell \) and \( \tau_{\ell} \leq n \).

It follows that

\[
\mathbb{P}(\tau_{\ell} < k) \leq e^{-q\ell - (qs+q^2\sigma^2/2)k} \quad q \geq 0.
\]

Minimizing the right-hand side of (72) over \( q \geq 0 \) gives

\[
\mathbb{P}(\tau_{\ell} < k) \leq e^{-(\ell-sk)^2/2\sigma^2k},
\]

which is obtained for \( q = q(k) = (\ell - sk)/\sigma^2k \). Note that this bound is valid for \( k \leq \ell/s \) since \( q \) should be nonnegative. By assumption \( k > 0 \), so inequality (10) follows from (73) by letting \( k = \ell/s - z \), \( 0 \leq z < \ell/s \).

Inequality (11) follows from Chernoff bound.
Claim ii. Using Claim i. and letting $u = \ell/s$, we have

$$
\mathbb{E} |\tau_\ell - \ell/s|^p = \int_0^\infty \mathbb{P}(|\tau_\ell - u| \geq z) d(z^p)
\leq 2 \int_0^\infty e^{-s^2 z^2/(2\sigma^2(u+z))} d(z^p)
\leq 2 \int_0^\infty e^{-s^2 z^2/(4\sigma^2 \max(u,z))} d(z^p)
= 2(1 + I_2)
$$

where the first inequality follows from Claim i. and where

$$
I_1 = \int_0^u e^{-s^2 z^2/(4\sigma^2 \ell)} d(z^p),
I_2 = \int_u^\infty e^{-s^2 z^2/(4\sigma^2 \ell)} d(z^p).
$$

For $I_1$, the change of variable

$$
z = \left(\frac{4\sigma^2 \ell}{s^3}\right)^{1/2} v^{1/p},
$$
yields

$$
I_1 = \left(\frac{4\sigma^2 \ell}{s^3}\right)^{p/2} \int_0^\infty e^{-v^{2/p}} dv
\leq \left(\frac{4\sigma^2 \ell}{s^3}\right)^{p/2} \int_0^\infty e^{-v^{2/p}} dv
= k_1 p/2 p > 0.
$$

where $0 < k_1 < \infty$ is a constant that depends on $s$, $p$, and $\sigma^2$.

For $I_2$, the change of variables $z = v^{1/p}$ and $v = ts^{-2p}$ yield

$$
I_2 = \int_0^\infty e^{-s^2 u^{1/p}/(4\sigma^2 \ell)} du
\leq e^{-s^2 u/(8\sigma^2)} \int_0^\infty e^{-s^2 v^{1/p}/(8\sigma^2)} dv
= e^{-s^2 u/(8\sigma^2)} s^{-2p} \int_0^\infty e^{-t^{1/p}/(8\sigma^2)} dt
\leq e^{-s^2 u/(8\sigma^2)} s^{-2p} \int_0^\infty e^{-t^{1/p}/(8\sigma^2)} dt
\leq s^{-2p} \int_0^\infty e^{-t^{1/p}/(8\sigma^2)} dt
= k_2
$$

where $0 < k_2 < \infty$ is a constant that depends on $s$, $p$, and $\sigma^2$. From (74), (75), and (76)

$$
\mathbb{E} |\tau_\ell - \ell/s|^p \leq k_3(k_4 + \ell^{p/2})
$$

for some constants $k_3$ and $k_4$ that depend on $s$, $p$, and $\sigma^2$. This yields the desired result.

Claim iii: See [5, Theorem 2.5].

Acknowledgments

The authors are grateful to the reviewers and to the Associate Editor for their insightful and detailed comments on the manuscript, and for questioning the non-asymptotic behavior of $\eta^*$ which prompted the investigation of the complementary stopping rule $\eta^{\ast}$.

The authors are also indebted to Milad Sefidgaran for many helpful discussions.

Appendix

Simulation - noisy observations: To numerically evaluate (7) for $\eta = \{\eta^\circ, \eta^\ast, \ell/s\}$, for each given value of $\ell$ we generated $n$ samples of $(X, Y)$, and computed the corresponding empirical sums

$$
s_n = \frac{1}{n} \sum_{i=1}^n |\eta(i) - \tau_\ell(i)| \quad \eta \in \{\eta^\circ, \eta^\ast, \ell/s\}
$$

where $(\eta(i), \tau_\ell(i))$ is the value of $(\eta, \tau_\ell)$ for the $i$-th sample of $(X, Y)$. To do this, we generated $n$ samples of $(X, Y)$, until

$$
C_1(\ell, p) = C_1(\ell, \varepsilon, s, p)
$$

be the constant defined in Theorem 2 with $\varepsilon = .5$ and $s = 10$. Chebyshev’s inequality gives the sufficient condition on the number of samples $n$

$$
n \geq \frac{\var{\eta - \tau_\ell}}{\delta^3 \cdot C_1^2(\ell, 1)}
$$

in order to have

$$
\mathbb{P} \left( \frac{1}{C_1(\ell, 1)} |s_n - \mathbb{E}|\eta - \tau_\ell| \right) \leq \delta \geq 1 - \delta.
$$

To use (77), we need to evaluate $\var{\eta - \tau_\ell}$. To do this, observe that $\mathbb{E}\eta \approx \mathbb{E}\tau_\ell \approx \ell/s$ for $\eta \in \{\eta^\circ, \eta^\ast, \ell/s\}$ (these approximations become equalities if we ignore overshoot). So we have

$$
\var{\eta - \tau_\ell} \approx \mathbb{E}|\eta - \tau_\ell|^2
= \left\{ \begin{array}{ll}
(1 + o(1))C_1(\ell, 2) & \eta = \eta^\circ \text{ or } \eta = \eta^\ast \\
(1 + o(1))(\ell/s)^3 & \eta = \ell/s
\end{array} \right.
$$

where the equality follows from Theorem 2 and (13). Combining (77) together with (79) gives

$$
n \geq \frac{\pi}{2 \cdot \delta^3} \quad \text{for } \eta = \eta^\circ \text{ or } \eta = \eta^\ast
n \geq \frac{5 \pi}{2 \cdot \delta^3} \quad \text{for } \eta = \ell/s
$$

11To be precise, we sequentially generated $(X_1, Y_1), (X_2, Y_2), \ldots$ until both $\tau_\ell$ and $\eta$ had stopped. So the generated samples $(X, Y)$’s are of variable length.
as a reasonable condition on \( n \) for (78) to hold. In Fig. 1
\( n = 10,000 \) which guarantees roughly \( \eta = \eta_0^{.05} \) or
\( \eta = \eta_{\ell}^{.05} \) and \( \delta = .1 \) for \( \eta = \ell / s \).

Finally note that, for small values of \( \ell \), the contribution due to
overshoot cannot be neglected and Theorem 2 is loose. So
in this regime the bounds (80) must be taken with a grain of
salt.

Simulation - delayed observations: We proceeded similarly
as in the previous section. We generated \( n \) samples \( X \),
computed the corresponding empirical sums \( s_n \) with \( \eta = \eta^* \),
and finally used Chebyshev’s related inequality (77) with
\( \text{Var}(\eta - \tau_2) = C_2(d, s, 2) \) and \( C_1(\ell, 1) \) replaced by \( C_2(d, s, 1) \)
to obtain
\[
\eta \geq \frac{C_2(d, s, 2)}{\ell^2} \cdot \frac{C_2(\ell, s, 1)}{\ell s} = \frac{\pi}{2 \delta^3}
\]  
(81)
as a reasonable condition on \( n \) to achieve \( \delta \) precision. In Fig. 2
\( n = 100,000 \) which guarantees a precision of \( \delta = .03 \).

BIographies

Marat V. Burnashev was born in Tashkent, USSR, on
January 6, 1947. He graduated and received Ph.D. degree
in mathematics from the Moscow Institute of Physics and
Technology in 1971 and 1974, respectively. He received the
Full Doctor degree in mathematics from the Mathematical
(Institute of the USSR Academy of Sciences in 1985.

Since 1974, he has been with the Institute for Information
Transmission Problems of the Russian Academy of Sciences.
He has also held visiting appointments at several universities
in the USA, Canada, Japan, Germany, France, etc.

His research interests include information theory, probabili-
ty theory, and mathematical statistics.

Aslan Tchamkerten received the Engineer Physicist Diploma in 2000 and the Ph.D. degree in Communications
in 2005, both from the Ecole Polytechnique Fédérale de
Lausanne (EPFL), Switzerland. Between 2005 and 2008, he
was a Postdoctoral Associate in the Department of Electrical
Engineering and Computer Science, Massachusetts Institute
of Technology (MIT), Cambridge. In 2008 he joined
Telecom ParisTech (ex. Ecole Nationale Supérieure des
Télécommunications, ENST), Paris, France, where he is cur-
rently Associate Professor. In 2009, he won a junior excellence
chair grant from the French National Research Agency (ANR).
His research interests are in Information Theory, Applied
Statistics, and Algorithms.

REFERENCES

[1] M. Basseville and I. Nikiforov. Detection of abrupt changes: theory and
application. Prentice-Hall, 1993.

[2] M. V. Burnashev and A. Tchamkerten. Tracking a Gaussian random
walk first-passage time through noisy observations. accepted in Ann.
App. Prob., 2010.

[3] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. Introduction
to Algorithms, 2nd edition. MIT Press, McGraw-Hill Book Company,
2000.

[4] M. H. DeGroot. Optimal Statistical Decisions. Wiley, Hoboken (NJ),
2004.

[5] A. Gut. On the moments and limit distributions of some first passage
times. Ann. Prob., 2(2):277–308, 1974.

[6] T.Z. Lai. Information bounds and quick detection of parameter changes
in stochastic systems. IEEE Trans. Inform. Th., 44:2917–2929, November
1998.

[7] A.A. Mogulskii. Absolute estimates for moments of certain boundary
functionals. Th. Prob. Appl., 18(2):350–357, 1973.

[8] G.V. Moustakides. Sequential change detection revisited. Ann. Statist.,
36(2):787–807, 1988.

[9] U. Niesen and A. Tchamkerten. Tracking stopping times through noisy
observations. IEEE Trans. Inform. Th., 55(1):422–432, January 2009.

[10] H.V. Poor and O. Hadjiliadis. Quickest detection. Cambridge, New
York, 2009.

[11] A. N. Shiryayev. On optimum methods in quickest detection problems.
Th. Prob. and its App., 8(1):22–46, 1963.

[12] A. N. Shiryayev. Optimal Stopping rules. Springer-Verlag, 1978.

[13] A. G. Tartakovsky and G. Moustakides. State-of-the-art in bayesian
changepoint detection. Seq. Analysis, 29(2):125–145, 2010.

[14] B. Yakir. Optimal detection of a change in distribution when the
observations form a Markov chain with a finite state space. In Change-
point problems, volume 23, pages 346–358. Institute of Mathematical
Statistics, Lecture Notes, Monograph Series, 1994.