Asymptotically Optimal Stochastic Encryption for Quantized Sequential Detection in the Presence of Eavesdroppers

Jiangfan Zhang®, Member, IEEE, and Xiaodong Wang®, Fellow, IEEE

Abstract—We consider sequential detection based on quantized data in the presence of eavesdropper. Stochastic encryption is employed as a counter measure that flips the quantization bits at each sensor according to certain probabilities, and the flipping probabilities are only known to the legitimate fusion center (LFC) but not the eavesdropping fusion center (EFC). As a result, the LFC employs the optimal sequential probability ratio test (SPRT) for sequential detection whereas the EFC employs a mismatched SPRT (MSPRT). We characterize the asymptotic performance of the MSPRT in terms of the expected sample size as a function of the vanishing error probabilities. We show that when the detection error probabilities are set to be the same at the LFC and EFC, every symmetric stochastic encryption is ineffective in the sense that it leads to the same expected sample size at the LFC and EFC. Next, in the asymptotic regime of small detection error probabilities, we show that every stochastic encryption degrades the performance of the quantized sequential detection at the LFC by increasing the expected sample size, and the expected sample size required at the EFC is no fewer than that is required at the LFC. Then the optimal stochastic encryption is investigated in the sense of maximizing the difference between the expected sample sizes required at the EFC and LFC. Although this optimization problem is nonconvex, we show that if the acceptable tolerance of the increase in the expected sample size at the LFC induced by the stochastic encryption is small enough, then the globally optimal stochastic encryption can be analytically obtained; and moreover, the optimal scheme only flips one type of quantized bits (i.e., 1 or 0) and keeps the other type unchanged.

Index Terms—Stochastic encryption, quantized sequential detection, mismatched SPRT, stopping time, eavesdropper, sensor networks.

I. INTRODUCTION

Decentralized detection in sensor networks using quantized data has been extensively studied, see [1]–[12] and references therein. The focus of this paper is on the sequential hypothesis testing in a decentralized sensor network based on quantized sensor data. Specifically, each sensor sequentially takes samples and then sends the binary quantized version of each sample to a legitimate fusion center (LFC). The LFC performs the sequential probability ratio test (SPRT) which is the optimal procedure for sequential detection of binary hypotheses in the sense of minimizing the expected sample sizes required for achieving the prescribed detection accuracy [13], [14].

Due to the broadcast nature of the communication links, the communications between sensors and the LFC are inherently vulnerable to eavesdropping, and hence, sensor networks are susceptible to security breach, which is an important problem especially when the network is part of a larger cyber-physical system [15]–[17]. For instance, some nodes within a cognitive radio (CR) network may eavesdrop on the transmissions from other nodes to the LFC, detect the vacant primary user channels, and encroach on the vacant primary user channels without paying any participation costs to the network moderator [16]. Eavesdropping fusion centers (EFC) in sensor networks are generally modeled as unauthorized receivers that passively wiretap communications between sensors and the LFC, have unbounded computational power just like the LFC, and seek to compete against the LFC [17].

Several security protocols have been recently proposed for sensor networks to defend against eavesdroppers [18], [19]. Most of these schemes employ the traditional network security protocols using authentication, cryptography, and key management techniques to provide security on the data link layer or the network layer. However, as pointed out in [20], the issue of scalability remains since these approaches provide security at the expense of increased bandwidth and energy consumption. Typically, large-scale sensor networks are comprised of low-cost sensor nodes with limited battery power, limited memory size, and low computing capacity. Due to the scarcity of resources in sensor networks, the security solutions summarized in [18], [19] which demand excessive processing, storage or communication overhead are not amenable to implementation in sensor networks. For example, public key ciphers are not suitable for sensor networks due to the high computational complexity. In this paper, as a low-complexity physical-layer security technique, stochastic encryption [16], [17], [21]–[23] is considered to be employed at the sensors such that every quantized data is transformed
according to certain probabilistic rule before transmitted to the LFC. The probabilistic transformation rule can be assigned to the sensors and the LFC prior to the deployment. As such, we can assume that the LFC has the prior knowledge of the probabilistic transformation rule, and hence, there is no need for the LFC and the sensors to exchange the information of the probabilistic transformation rule after the deployment of the sensors. The EFC is assumed to have less information about the sensors than the LFC, and it is assumed to be ignorant of the information of the probabilistic transformation rule. Due to the optimality of the SPRT, it is natural for the EFC to also employ the SPRT but based on its own prior knowledge to make its decision. Otherwise, the LFC can employ the SPRT to easily defeat the EFC even without using any defending approach. More detailed justification about the assumption on the EFC’s sequential detection strategy will be provided in Section IV-D.

In the context of sequential detection, the LFC sequentially receives a stream of observations, and needs to check the stopping condition every time it receives an observation. In light of this, the encryption and decryption of each observation should not depend on its subsequent observations to ensure the timely sequential implementation of the detection process at the LFC. If the encryption and decryption of each observation depend on its preceding observations, then each sensor and the LFC may need additional memory to store the corresponding preceding observations, which brings about a dispensable increase in the cost of systems and hence should be forestalled. To this end, the encryption and decryption of each observation should not depend on its subsequent and preceding observations in such sequential detection systems. As such, when compared with other cryptographic algorithms, employing stochastic encryptions is an easy way to provide physical layer security for sequential detection with quantized data, since it does not introduce any communication overhead for the sensors and has minimal processing requirements, rendering it scalable in terms of network size. For example, considering the RSA algorithm [24], which is one of the first public-key cryptosystems and is widely used for secure data transmission, the typical key size ranges from 1024 bits to 4096 bits, and the size of the encrypted data produced by the RSA algorithm is the same as the key size. Hence, the RSA algorithm may introduce huge communication overhead especially when each unencrypted quantized observation only has few bits. To be specific, if each unencrypted quantized observation is just one-bit, which is the case that we are primarily interested in, then the size of the corresponding encrypted observation is 1024-bits when the RSA algorithm with 1024-bits key size is employed, and therefore, additional 1023 bits are required to sent from the sensor to the LFC to secure the transmission of every unencrypted observation. On the other hand, if the stochastic encryption is employed, then the size of the encrypted observation is the same as that of the unencrypted observation, which implies that the stochastic encryption does not introduce any communication overhead, and hence is more desirable than the RSA algorithm from the perspective of communication efficiency.

Since the stochastic encryption just simply flips each quantized data with some probability, the complexity of the computations required by the stochastic encryption is very low as long as each sensor has some source for generating random numbers. This can be easily addressed by equipping each sensor with a hardware random number generator which is generally inexpensive and widely used in many other cryptographic algorithms and protocols, such as the aforementioned RSA algorithm [24], [25]. Alternatively, each sensor can resort to a pseudo-random number generator which is commonly implemented by means of algorithms. On the other hand, if the RSA algorithm is employed, then each sensor also requires some source for generating random numbers, and moreover, the computations required by the RSA algorithm, such as modular exponentiation, are much more complicated than that required by the stochastic encryption [24]. In addition, unlike many other cryptographic algorithms and protocols, if the stochastic encryption is employed, the LFC does not need to decrypt the received encrypted observations, and can just use the encrypted observations to implement the sequential detection process. Hence, employing the stochastic encryption can dramatically reduce the complexity of the computations required at the receiving end of the system.

It is worth mentioning that it is possible to employ the stochastic encryption, which is a physical layer security technique, in conjunction with other security techniques on higher layers, such as the security techniques on the data link layer and the network layer, to enhance the integrity of sensor network operation. Admittedly, the stochastic encryption provides security at the cost of inducing local randomness at the sensors, which may degrade the performance of the LFC. However, it will be shown in this paper that the randomness induced at the sensors can greatly confuse the EFC, and hence can effectively help the LFC beat the EFC. In this paper, we first investigate the sequential detection performance of the EFC in the presence of stochastic encryption, and then optimize the encryption scheme to maximize the difference between the expected sample sizes at the EFC and LFC.

A. Summary of Results and Main Contributions

Since the EFC is unaware of the stochastic encryption parameters, a mismatched SPRT (MSPRT) is employed at the EFC. We characterize the expected sample size and the error probabilities of the MSPRT in terms of the detection thresholds. We show that when the detection error probabilities are set to be the same at the LFC and EFC, every symmetric stochastic encryption leads to the same expected sample size at the LFC and EFC. In addition, the asymptotic analysis on the expected sample size in terms of the vanishing error probabilities is provided, and the stark difference from the asymptotic performance of the SPRT with no model mismatch is revealed. For example, the expected sample size of the SPRT is determined by the Kullback-Leibler (KL) divergences between the distributions under the two hypotheses, while the expected sample size of the MSPRT is unrelated to the KL divergences.
Next, in the asymptotic regime of small error probabilities, we show that every stochastic encryption degrades the performance of the quantized sequential detection at the LFC by increasing the expected sample size, and the expected sample size required at the EFC is no fewer than that is required at the LFC. Hence, symmetric stochastic encryptions are the least effective ones. Then the optimal stochastic encryption is investigated in the sense of maximizing the difference between the expected sample sizes required at the EFC and LFC. The optimization problem is nonconvex. However we show that if the acceptable tolerance of the increase in the expected sample size at the LFC induced by the stochastic encryption is small enough, then the globally optimal stochastic encryption can be analytically obtained. Moreover, the optimal scheme only flips one type of quantized bits (i.e., 1 or 0) that has larger probability and keeps the other type unchanged.

A. Quantized Sequential Detection

Consider a sensor network consisting of an LFC and $N$ spatially distributed sensors, which aims to test between two hypotheses. Each sensor sequentially makes observations of a particular phenomenon. Let $x_k^{(n)}$ denote the $k$-th observation made at the $n$-th sensor. Under each hypothesis, the observations are assumed to be independent and identically distributed (i.i.d.) at each sensor and are independent across sensors. We use $f_{0}^{(n)}(x)$ and $f_{1}^{(n)}(x)$ to denote the probability density functions (pdf) of the observations under the two hypotheses, i.e., for all $n \in \{1, 2, ..., N\}$ and for all $k \in \{1, 2, ..., N\}$

$$
\mathcal{H}_0 : x_k^{(n)} \sim f_{0}^{(n)}(x),
$$

and

$$
\mathcal{H}_1 : x_k^{(n)} \sim f_{1}^{(n)}(x).
$$

(1)

For each observation $x_k^{(n)}$, the $n$-th sensor first forms a one-bit summary message $u_k^{(n)}$ by applying a quantizer $Q_n(x)$ with the quantization region $D_n$, i.e.,

$$
u_k^{(n)} \triangleq Q_n(x_k^{(n)}) \in \{0, 1\},
$$

(2)

and then sends $u_k^{(n)}$ to the LFC.

It is clear that $\{u_k^{(n)}\}$ are independent and identically distributed at each sensor and are independent across sensors. Let $p_n$ and $q_n$ denote the probabilities of the event $\{u_k^{(n)} = 1\}$ under the hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively, i.e.

$$
p_n \triangleq \int_{x \in D_n} f_{1}^{(n)}(x) \, dx \quad \text{and} \quad q_n \triangleq \int_{x \in D_n} f_{0}^{(n)}(x) \, dx.
$$

(3)

Then the log-likelihood ratio (LLR) of the quantized data $u_k^{(n)}$ can be computed as

$$
l_k^{(n)} = \begin{cases} 
\ln \frac{p_n}{q_n} & \text{if } u_k^{(n)} = 1, \\
\ln \frac{1 - p_n}{1 - q_n} & \text{if } u_k^{(n)} = 0.
\end{cases}
$$

(4)

With regard to the probabilities $\{p_n\}$ and $\{q_n\}$, the following assumption is made throughout the paper.

**Assumption 1.** We assume that the local quantizers $\{Q_n\}_{n=1}^{N}$ bring about $p_n = p \in (0.5, 1)$, $q_n = q \in (0, 0.5)$, and $p + q = 1$ for all $n \in \{1, 2, ..., N\}$.

It is worth mentioning that Assumption 1 is motivated by the classical problem of detecting the mean shift in Gaussian noise [8], [30]. Specifically, assume the following model and equal priors on both hypotheses

$$
\mathcal{H}_0 : x_k^{(n)} = w_k^{(n)},
$$

and

$$
\mathcal{H}_1 : x_k^{(n)} = \theta + w_k^{(n)},
$$

(5)

where $\theta$ is a deterministic quantity, and the independent Gaussian noise $w_k^{(n)} \sim N(\sigma^2)$. As claimed in [8], in the sense of minimizing the expected sample size at the LFC, the one-bit optimal threshold quantizer at each sensor is symmetric, i.e., $D_n = \{x_k^{(n)} : x_k^{(n)} \geq \alpha + \frac{\theta}{2}\}$, if a prescribed upper bound on the false alarm and miss probabilities is given. By employing this quantizer at each sensor, it is easy to show

B. Related Works

The decentralized sequential detection in sensor networks using quantized data has been widely investigated, see [4]–[12] for instance. However, to the best of our knowledge, stochastic encryption for quantized sequential detection in the presence of eavesdroppers has not been considered.

Stochastic encryptions were originally proposed in [21] for physical-layer security in the context of fixed-sample-size estimation problems with quantized data. For the fixed-sample-size hypothesis testing in sensor networks, the joint design of the stochastic encryption and the LFC decision rule that minimizes the LFC detection error probabilities subject to a constraint on the EFC error probabilities is studied in [22]. Nonetheless, the design approach in [22] is ad hoc and results in a suboptimal stochastic encryption. This design approach is made more rigorous in [23], where the optimal stochastic encryption is obtained with respect to the J-divergence which is made more rigorous in [23], where the optimal stochastic encryption is obtained with respect to the J-divergence. This design approach is made more rigorous in [23], where the optimal stochastic encryption is obtained with respect to the J-divergence. This design approach is made more rigorous in [23], where the optimal stochastic encryption is obtained with respect to the J-divergence.

From the EFC perspective, the stochastic encryption process can be treated as a malicious man-in-the-middle attack [26]–[29], since the EFC is unaware of the probabilistic transformation of the quantized sensor data. However, most existing works on the man-in-the-middle attacks focus on the fixed-sample-size inference problems, and do not jointly consider the performance at the LFC and EFC.

The remainder of the paper is organized as follows. The system model and stochastic encryption is introduced in Section II. The performance of the mismatched SPRT employed by the EFC is analyzed in Section III-C. In Section IV, the optimal stochastic encryption is pursued. Finally, Section V provides our conclusions.

II. SYSTEM MODEL AND STOCHASTIC ENCRYPTION

In this section, the system model and the stochastic encryption are introduced. The sequential decision procedures adopted at the LFC and EFC are also specified.
that all conditions in Assumption 1 are satisfied. As shown later, Assumption 1 can facilitate our analysis.

We assume that the LFC can reliably receive the quantized data from the sensors. While sequentially receiving the quantized data from the sensors, the LFC implements a sequential decision procedure to test between the hypotheses in (1). Besides the LFC, there exists an EFC in the sensor system which is able to wiretap the quantized sensor data transmitted from the sensors to the LFC, and the EFC also aims to perform sequential detection between the hypotheses in (1). Our goal is to design a strategy at the sensors to transform the quantized data \( \{ u_k^{(n)} \} \) so that under the same detection performance constraint, the LFC will reach the decision faster than the EFC since the former is aware of the transformation but the latter is not.

B. Stochastic Encryption and the SPRT at the LFC

The idea is to stochastically encrypt the quantized data \( \{ u_k^{(n)} \} \) at each sensor before they are transmitted to the LFC. To be specific, at each sensor, an encrypted version \( \tilde{u}_k^{(n)} \) of the quantized data \( u_k^{(n)} \) is reported to the LFC which follows the encryption rule

\[
\begin{aligned}
\mathbb{P} \left( \tilde{u}_k^{(n)} = 1 \mid u_k^{(n)} = 0 \right) &= \psi_1, \\
\mathbb{P} \left( \tilde{u}_k^{(n)} = 0 \mid u_k^{(n)} = 1 \right) &= \psi_0,
\end{aligned}
\]

where \( \psi_0, \psi_1 \in [0, 1] \). Hence the quantized data \( u_k^{(n)} \) is flipped with probability \( \psi_i \) if \( u_k^{(n)} = i \) for \( i \in \{0, 1\} \). Let \( \tilde{p} \) and \( \tilde{q} \) denote the probabilities of the event \( \{ \tilde{u}_k^{(n)} = 1 \} \) under the hypotheses \( \mathcal{H}_1 \) and \( \mathcal{H}_0 \), respectively. Then we can obtain

\[
\tilde{p} = (1 - \psi_0 - \psi_1) p + \psi_0, \quad \tilde{q} = (1 - \psi_0 - \psi_1) q + \psi_0.
\]

We assume that the LFC is aware of the encryption parameters \( \psi_0 \) and \( \psi_1 \), while the EFC does not know the existence of the stochastic encryptions.

Under Assumption 1, every encrypted bit received at the LFC is independent and follows the same distribution. Hence, for notational simplicity, we use \( \tilde{u}_t \) to denote the \( t \)-th encrypted bit received at the LFC henceforth. In general, the sequential detection procedure employed by the LFC consists of a stopping rule \( T_L \) and a decision function \( D_L \). The stopping rule \( T_L \) specifies when the sequential test stops for decision, and upon stopping at \( T_L \), the decision function \( D_L \) chooses between the two hypotheses. The optimal sequential detector (\( T_L, D_L \)) minimizes the expected number of sensor data required to reach a decision with probabilities of false alarm and miss upper bounded by \( \alpha_L^* \) and \( \beta_L^* \), respectively. It is well known that Wald’s sequential probability ratio test (SPRT) achieves this optimality [14]. Thus, we assume that the LFC employs the SPRT with the test statistic

\[
\tilde{L}_t \triangleq \sum_{s=1}^{t} \tilde{i}_s,
\]

where \( \tilde{i}_s \) denotes the LLR of the \( s \)-th received encrypted bit, i.e.,

\[
\tilde{i}_s = \mathbb{I}_{\{ \tilde{u}_s = 1 \}} \ln \frac{\tilde{p}}{\tilde{q}} + \mathbb{I}_{\{ \tilde{u}_s = 0 \}} \ln \frac{1 - \tilde{p}}{1 - \tilde{q}}.
\]

The stopping rule and the decision function are given respectively by

\[
T_L \triangleq \inf \left\{ t \mid \tilde{L}_t \notin (-A_L, B_L) \right\},
\]

\[
D_L \triangleq \begin{cases} 1, & \text{if } \tilde{L}_{T_L} \geq B_L, \\ 0, & \text{if } \tilde{L}_{T_L} \leq -A_L, \end{cases}
\]

where the thresholds \( A_L \) and \( B_L \) are chosen such that \( \mathbb{P}_0 (D_L = 1) = \alpha_L^* \) and \( \mathbb{P}_1 (D_L = 0) = \beta_L^* \). The SPRT given by (11)–(12) is optimal in the sense of minimizing both \( \mathbb{E}_0 \{ T_L \} \) and \( \mathbb{E}_1 \{ T_L \} \), where \( \mathbb{P}_i \) and \( \mathbb{E}_i \{ \cdot \} \) denote the probability measure and the expectation operator under \( \mathcal{H}_i \), respectively.

From (10), the LLR is bounded from above as per

\[
\tilde{L}_t \leq \max \left\{ \ln \frac{\tilde{p}}{\tilde{q}}, \ln \frac{1 - \tilde{p}}{1 - \tilde{q}} \right\},
\]

which yields the following result.

**Proposition 1.** As \( \alpha_L^* \) and \( \beta_L^* \) tend to 0 in such a way that \( \alpha_L^* + \beta_L^* < 1 \), \( \alpha_L^* \ln \beta_L^* \rightarrow 0 \) and \( \beta_L^* \ln \alpha_L^* \rightarrow 0 \), the asymptotic performance of the SPRT employed at the LFC is characterized as

\[
\mathbb{E}_0 \{ T_L \} = -\frac{\alpha_L^* \ln \frac{\alpha_L^*}{1 - \alpha_L^*} + (1 - \alpha_L^*) \ln \frac{1 - \alpha_L^*}{\beta_L^*}}{\tilde{q} \ln \frac{\tilde{p}}{\tilde{q}} + (1 - \tilde{q}) \ln \frac{1 - \tilde{q}}{1 - \tilde{p}}} + O(1),
\]

\[
\mathbb{E}_1 \{ T_L \} = -\frac{\beta_L^* \ln \frac{\beta_L^*}{1 - \beta_L^*} + (1 - \beta_L^*) \ln \frac{1 - \beta_L^*}{\alpha_L^*}}{\tilde{p} \ln \frac{\tilde{q}}{\tilde{p}} + (1 - \tilde{p}) \ln \frac{1 - \tilde{p}}{1 - \tilde{q}}} + O(1),
\]

The proof of Proposition 1 is omitted here, since it is similar to the proof of Theorem 3.1.4 in [31]. It is worth mentioning that as \( \alpha_L^* \) and \( \beta_L^* \) go to 0, \( M_L^{(0)} \) and \( M_L^{(1)} \) in (13) and (14) respectively increase to infinity and dominate the \( O(1) \) terms, and therefore, determine the behavior of \( \mathbb{E}_0 \{ T_L \} \) and \( \mathbb{E}_1 \{ T_L \} \), respectively.

III. MISMATCHED SPRT AT THE EFC

In this section, we first describe the EFC model, and show that symmetric stochastic encryptions are ineffective, since the expected sample sizes at the LFC and EFC are identical when the detection error probabilities are the same at the LFC and EFC. Then, we obtain the explicit asymptotic characterization of the expected sample size of the MSPRT.

A. Eavesdropping Fusion Center Model

We assume that the EFC can reliably receive all the quantized data transmitted from the sensors to the LFC, and the EFC is aware of \( f_0^{(n)}(x) \), \( f_1^{(n)}(x) \) and the quantizers employed at the sensors. Thus, the EFC can figure out \( p \) and \( q \) by using (3). However, we assume that the EFC is unaware
of $\psi_0$ and $\psi_1$ which essentially are the keys of the stochastic encryption.

Based on this EFC model, we assume that the EFC implements the SPRT but based on a mismatched model, since it is unaware of the stochastic encryption parameters, i.e., $\psi_0$ and $\psi_1$. We refer to such sequential detection procedure as the mismatched SPRT (MSPRT). More detailed justification for the assumption that the EFC employs the MSPRT will be provided in Section IV-D.

B. Mismatched SPRT and Ineffective Stochastic Encryptions

From (4), the mismatched log-likelihood ratio of the $s$-th encrypted bit $\tilde{u}_s$, which is based on the unencrypted data model, can be written as

$$l_s = \sum_{s=1}^{t} l_s = \sum_{s=1}^{t} \eta \left(1_{\{\tilde{u}_s=1\}} - 1_{\{\tilde{u}_s=0\}}\right),$$

(16)

where $\eta \triangleq \ln \frac{p}{q}$. By comparing (16) with (20), we can obtain

$$\tilde{L}_t = \frac{\tilde{\eta}}{\eta} L_t,$$

(21)

Noting that $A_E$ and $B_E$ are chosen to be multiples of $\eta$, no overshoot effect occurs in MSPRT. Hence, we can obtain

$$\alpha_E = \mathbb{P}_0 (D_E = 1) = \mathbb{E}_0 \left\{1 \{L_{TE}=B_E\}\right\}$$

$$= \mathbb{E}_1 \left\{e^{-L_{TE}} 1 \{L_{TE}=B_E\}\right\} = e^{-\frac{B_E}{\eta}} (1 - B_E),$$

(22)

which implies

$$B_E = \frac{\eta}{\alpha_E} \ln \frac{1 - B_E}{\beta_E}.$$

(23)

On the other hand, from (11) and (20), the stopping rule $T_L$ can be simplified to

$$T_L = \inf \left\{t \sum_{s=1}^{t} (1_{\{\tilde{u}_s=1\}} - 1_{\{\tilde{u}_s=0\}}) \not\in \left(-A_L, \frac{B_L}{\eta}\right) \right\},$$

(24)

and hence, $A_L$ and $B_L$ can be chosen to be multiples of $\tilde{\eta}$ so that no overshoot effect occurs. From the definition of the false alarm probability, we can obtain

$$\alpha_L = \mathbb{P}_0 (D_L = 1) = \mathbb{E}_0 \left\{1 \{L_{TL}=B_L\}\right\}$$

$$= \mathbb{E}_1 \left\{e^{-L_{TL}} 1 \{L_{TL}=B_L\}\right\} = e^{-B_L} (1 - \beta_L),$$

(25)

which yields

$$B_L = \ln \frac{1 - \beta_L}{\alpha_L}.$$

(26)

It is seen from (23) and (26) that if $\alpha_L = \alpha_E$ and $\beta_L = \beta_E$, then

$$\frac{B_E}{\eta} = \frac{B_L}{\tilde{\eta}}.$$

(27)

Similarly, by employing the definition of the miss probability, we can obtain

$$\frac{A_E}{\eta} = \frac{A_L}{\tilde{\eta}}.$$

(28)

As a result, by comparing (17) with (24), we know that if $\alpha_L = \alpha_E$ and $\beta_L = \beta_E$, then

$$T_L = T_E,$$

(29)

which concludes the proof.

The expression of $B_E$ in (23) deserves some discussion. Since the EFC is unaware of the stochastic encryption, the EFC, in practice, designs the thresholds $A_E$ and $B_E$ based on the prescribed detection error constraint and the mismatched model which does not depend on the encryption parameters $\psi_0$ and $\psi_1$. Therefore, the designed $A_E$ and $B_E$ should not depend on the encryption parameters $\psi_0$ and $\psi_1$ in practice. However, the actual detection error of the EFC may not satisfy the prescribed detection error constraint in practice.
due to the model mismatch. If the designed $A_E$ and $B_E$ at the EFC give rise to some detection error which is much larger than the prescribed detection error, then it is possible that the expected sample size required at the EFC is smaller than that required at the LFC. For a fair comparison between the expected sample sizes required at the EFC and that required at the LFC, we need to adopt the thresholds $A_E$ and $B_E$ which can satisfy the prescribed detection error constraint at the EFC. This is the reason why $B_E$ in (23) depends on the encryption parameters $\psi_0$ and $\psi_1$.

C. Expected Sample Size and Detection Performance of the Mismatched SPRT

In this subsection, we analyze the expected sample size and the detection performance of the MSPRT for a given pair of thresholds $(A_E, B_E)$.

In order to obtain $\mathbb{E}_0\{T_E\}$ for $i = 0, 1$, we will make use of a sequence of stopping times $\{T^{(m)}\}_{m \in \mathbb{Z}}$ which are defined as

$$T^{(m)} \triangleq \inf \left\{ t \in \mathbb{Z} \setminus \{0\} : t > m + \sum_{s=1}^{t} \left( \mathbb{I}_{\{\tilde{u}_s = 1\}} - \mathbb{I}_{\{\tilde{u}_s = 0\}} \right) \neq 0 \right\}. \quad (30)$$

It is clear that

$$T^{(m)} = 0,$$

if $m \leq \frac{A_E}{\eta}$ or $m \geq \frac{B_E}{\eta}$. \quad (31)

According to the definition of $T^{(m)}$ in (30) and the distribution of $\tilde{u}_s$, we can obtain that for any $m \in (-A_E/\eta, B_E/\eta)$,

$$\mathbb{E}_0\{T^{(m)}\} = \tilde{q}\mathbb{E}_0\{T^{(m+1)}\} + (1 - \tilde{q})\mathbb{E}_0\{T^{(m-1)}\} + 1. \quad (32)$$

Furthermore, the boundary condition in (31) implies

$$\mathbb{E}_0\{T^{(-A_E/\eta)}\} = \mathbb{E}_0\{T^{(B_E/\eta)}\} = 0. \quad (33)$$

Solving the recursion given by (32)–(33), we can obtain that if $\tilde{q} \neq \frac{1}{2}$, then

$$\mathbb{E}_0\{T^{(m)}\} = \left[ 1 - \left( \frac{1-q}{\tilde{q}} \right)^{m + \frac{A_E}{\eta}} \right] \frac{A_E + B_E}{\eta} - m \left( \frac{A_E + B_E}{\eta} \right) \left( \frac{1-q}{\tilde{q}} \right)^{m} \left( 1 - \left( \frac{1-q}{\tilde{q}} \right)^{2q} \right) \left( 2\tilde{q} - 1 \right), \quad (34)$$

and if $\tilde{q} = \frac{1}{2}$, then

$$\mathbb{E}_0\{T^{(m)}\} = \left( m + \frac{A_E}{\eta} \right) \left( B_E \frac{1-q}{\tilde{q}} - m \right). \quad (35)$$

By comparing (17) and (30), we can obtain that $T^{(0)} = T_E$, and hence, if $\tilde{q} \neq \frac{1}{2}$,

$$\mathbb{E}_0\{T_E\} = \left[ 1 - \left( \frac{1-q}{\tilde{q}} \right)^{A_E + B_E} \right] \frac{A_E + B_E}{\eta} \left( 2\tilde{q} - 1 \right), \quad (36)$$

and if $\tilde{q} = \frac{1}{2}$,

$$\mathbb{E}_0\{T_E\} = \frac{A_E B_E}{\eta^2}. \quad (37)$$

By employing a similar approach, we obtain that if $\tilde{q} \neq \frac{1}{2}$,

$$\mathbb{E}_1\{T_E\} = \left[ 1 - \left( \frac{1-q}{\bar{q}} \right)^{A_E + B_E} \right] \frac{A_E + B_E}{\eta} \left( 2\bar{q} - 1 \right), \quad (38)$$

and if $\tilde{q} = \frac{1}{2}$,

$$\mathbb{E}_1\{T_E\} = \frac{A_E B_E}{\eta^2}. \quad (39)$$

Next, we consider the detection performance of the MSPRT. Let $\alpha_E$ and $\beta_E$ denote the false alarm and miss probabilities of the MSPRT at the EFC, respectively, that is,

$$\alpha_E \triangleq \mathbb{P}_0\{L_{T_E} \geq B_E\} = \mathbb{P}_0\left( \sum_{s=1}^{T_E} (\mathbb{I}_{\{\tilde{u}_s = 1\}} - \mathbb{I}_{\{\tilde{u}_s = 0\}}) \geq \frac{B_E}{\eta} \right), \quad (40)$$

and $\beta_E \triangleq \mathbb{P}_1\{L_{T_E} \leq -A_E\} = \mathbb{P}_1\left( \sum_{s=1}^{T_E} (\mathbb{I}_{\{\tilde{u}_s = 1\}} - \mathbb{I}_{\{\tilde{u}_s = 0\}}) \leq -\frac{A_E}{\eta} \right). \quad (41)$

Based on the sequence of stopping times $\{T^{(m)}\}_{m \in \mathbb{Z}}$, we also define a sequence of probabilities $\{\alpha^{(m)}\}_{m \in \mathbb{Z}}$ as

$$\alpha^{(m)} \triangleq \mathbb{P}_0\left( m + \sum_{s=1}^{T^{(m)}} (\mathbb{I}_{\{\tilde{u}_s = 1\}} - \mathbb{I}_{\{\tilde{u}_s = 0\}}) \geq \frac{B_E}{\eta} \right), \quad (42)$$

which yields that $\forall m \in (-A_E/\eta, B_E/\eta)$

$$\alpha^{(m)} = \tilde{q}\alpha^{(m+1)} + (1 - \tilde{q})\alpha^{(m-1)}. \quad (43)$$

by employing the distribution of $\tilde{u}_s$. What is more, the boundary condition in (31) implies

$$\alpha^{(-A_E/\eta)} = 0 \quad \text{and} \quad \alpha^{(B_E/\eta)} = 1. \quad (44)$$

Solving the recursion in (43)–(44), we can obtain

$$\alpha^{(m)} = \begin{cases} \frac{1 - \left( \frac{1-q}{\tilde{q}} \right)^{m + \frac{A_E}{\eta}}}{1 - \left( \frac{1-q}{\tilde{q}} \right)^{A_E + B_E} \frac{A_E + B_E}{\eta}}, & \text{if } \tilde{q} \neq \frac{1}{2}, \\ \frac{1}{A_E + B_E \left( m + \frac{A_E}{\eta} \right)}, & \text{if } \tilde{q} = \frac{1}{2}. \end{cases} \quad (45)$$

The results in (36)–(39) can also be obtained by using the results on Gambler’s Ruin problem.
From (40) and (42), we know $\alpha_E = \alpha^{(0)}$, and therefore, $\alpha_E$ can be obtained from (45) as

$$\alpha_E = \begin{cases} 
1 - \left(1 - \frac{1}{\bar{q}}\right)^{\frac{\Delta_E}{\eta}} / \left(1 - \left(1 - \frac{1}{\bar{q}}\right)^{\frac{\Delta_E}{\eta}}\right), & \text{if } \bar{q} \neq \frac{1}{2}, \\
A_E / A_E + B_E, & \text{if } \bar{q} = \frac{1}{2}.
\end{cases}$$ (46)

Similarly, the miss probability of the MSPRT can be derived as

$$\beta_E = \begin{cases} 
1 - \left(1 - \frac{1}{\bar{p}}\right)^{\frac{\Delta_E}{\eta}} + \frac{\mu}{\nu} \left(\frac{\mu}{\nu} - \mu\right)^{\frac{\Delta_E}{\eta}}, & \text{if } \bar{p} \neq \frac{1}{2}, \\
B_E / A_E + B_E, & \text{if } \bar{p} = \frac{1}{2}.
\end{cases}$$ (47)

**D. Asymptotic Characterization of the Mismatched SPRT**

It is seen from (36)–(39), (46) and (47) that $E_i(T_E), \alpha_E$ and $\beta_E$ can all be exactly expressed in terms of the thresholds $A_E$ and $B_E$. Note that (46) and (47) are transcendental equations, and therefore, do not admit closed-form expressions of $A_E$ and $B_E$ in terms of $\alpha_E$ and $\beta_E$. Thus, it is generally difficult to express $E_i(T_E)$ in terms of $\alpha_E$ and $\beta_E$. In this subsection, we focus on the asymptotic regime where $\alpha_E$ and $\beta_E$ tend to zero, and derive the asymptotic approximations of $E_i(T_E)$ in terms of $\alpha_E$ and $\beta_E$.

It will be proved in Lemma 1 in Section IV-C that the larger the stochastic encryption parameters, the larger the expected detection delay at the LFC when the probabilities of false alarm and miss at the LFC are small. Thus, small stochastic encryption parameters are generally more desirable in many practical applications, for example attack detection in smart grid systems [32], since a prompt decision is critical for such problems. It is seen from (7) and (8) that when the stochastic encryption parameters $\psi_0$ and $\psi_1$ are small, $\bar{p}$ and $\bar{q}$ are close to $p$ and $q$, respectively. Moreover, $p$ and $q$ are larger and smaller than 0.5, respectively. In light of this, we make the following assumption.

**Assumption 2.** The stochastic encryption parameters $\psi_0$ and $\psi_1$ are small enough so that

$$\bar{p} = (1 - \psi_0 - \psi_1)p + \psi_0 \geq \frac{1}{2}$$ (48)

and

$$\bar{q} = (1 - \psi_0 - \psi_1)q + \psi_0 \leq \frac{1}{2}.\ (49)$$

The following theorem characterizes the asymptotic performance of the MSPRT at the EFC when $\alpha_E$ and $\beta_E$ tend to zero.

**Theorem 2.** Under Assumptions 1 and 2, the following results hold.

1) As $\alpha_E, \beta_E \to 0$, the thresholds $A_E, B_E \to \infty$.

2) The expected sample sizes $E_i(T_E)$ at the EFC under $\mathcal{H}_0$ and $\mathcal{H}_1$ can be respectively written in terms of $\alpha_E$ and $\beta_E$ as

$$\bar{E}_0 \{T_E\} = \frac{1}{1 - 2q} \left[ (1 - \alpha_E) \log \frac{1}{\beta_E} - \alpha_E \log \frac{1}{\alpha_E} \right]$$ (50)

$$\bar{E}_1 \{T_E\} = \frac{1}{2\bar{p} - 1} \left[ (1 - \beta_E) \log \frac{1}{\alpha_E} - \beta_E \log \frac{1}{\beta_E} \right]$$ (51)

with $\mu \triangleq \frac{\bar{p}}{1 - \bar{p}} > 1$ and $\nu \triangleq \frac{1 - \bar{q}}{\bar{q}} > 1$. (52)

3) For any given $\psi_0$ and $\psi_1$, we have

$$\frac{\partial M_E^{(0)}}{\partial \alpha_E} < 0 \text{ and } \frac{\partial M_E^{(1)}}{\partial \alpha_E} < 0,$$ (53)

$$\frac{\partial M_E^{(0)}}{\partial \beta_E} < 0 \text{ if } \alpha_E < \frac{1}{e},$$ (54)

$$\frac{\partial M_E^{(1)}}{\partial \beta_E} < 0 \text{ if } \beta_E < \frac{1}{e}.$$ (55)

**Proof:** We first prove 1). Under Assumptions 1 and 2, by employing (46), (47) and (52), the expressions of $\alpha_E$ and $\beta_E$ can be simplified to

$$\alpha_E = \frac{\nu^{\frac{\Delta_E}{\eta}} - 1}{\nu^{\frac{\Delta_E}{\eta}} - \nu^{-\frac{\Delta_E}{\eta}}} = \frac{1 - \nu^{-\frac{\Delta_E}{\eta}}}{\nu^{\frac{\Delta_E}{\eta}} - \nu^{-\frac{\Delta_E}{\eta}}},$$ (56)

and

$$\beta_E = \frac{\mu^{\frac{\Delta_E}{\eta}} - 1}{\mu^{\frac{\Delta_E}{\eta}} - \mu^{-\frac{\Delta_E}{\eta}}} = \frac{1 - \mu^{-\frac{\Delta_E}{\eta}}}{\mu^{\frac{\Delta_E}{\eta}} - \mu^{-\frac{\Delta_E}{\eta}}}.\ (57)$$

Since $A_E/\eta \geq 1$ and $B_E/\eta \geq 1$, $\alpha_E$ and $\beta_E$ can be bounded from below as per

$$\alpha_E > \frac{1 - \nu^{-1}}{\nu^{-\frac{\Delta_E}{\eta}}} \text{ and } \beta_E > \frac{1 - \mu^{-1}}{\mu^{-\frac{\Delta_E}{\eta}}}.$$ (58)

which implies that as $\alpha_E \to 0$, $B_E \to \infty$, and similarly, $A_E \to \infty$ as $\beta_E \to 0$.

Next, we prove 2). From (56), we can obtain

$$\alpha_E \nu^{-\frac{\Delta_E}{\eta}} - 1 = -(1 - \alpha_E) \nu^{-\frac{\Delta_E}{\eta}}.$$ (59)

Note that $A_E \to \infty$, as $\beta_E \to 0$, and hence from (59), we know

$$\alpha_E \nu^{-\frac{\Delta_E}{\eta}} - 1 = o(1), \text{ as } \beta_E \to 0,$$ (60)

which implies

$$\frac{B_E}{\eta} = \log_{\nu} \frac{1}{\alpha_E} + o(1), \text{ as } \beta_E \to 0.$$ (61)

Similarly, (57) yields that

$$\frac{A_E}{\eta} = \log_{\mu} \frac{1}{\beta_E} + o(1), \text{ as } \alpha_E \to 0.$$ (62)
By employing (36), (52), (56), (61) and (62), we can obtain that as \( \alpha_E, \beta_E \to 0 \)

\[
\mathbb{E}_0 \{ T_E \} = \frac{1}{1-2q} \left[ (1 - \alpha_E) \frac{A_E}{\eta} - \alpha_E \frac{B_E}{\eta} \right] = \frac{1}{1-2q} \left[ (1 - \alpha_E) \log_\nu \frac{1}{\beta_E} - \alpha_E \log_\nu \frac{1}{\alpha_E} \right] + o(1).
\]

Similarly, we can show that as \( \alpha_E, \beta_E \to 0 \),

\[
\mathbb{E}_1 \{ T_E \} = \frac{1}{2\hat{p} - 1} \left[ (1 - \beta_E) \log_\nu \frac{1}{\alpha_E} - \beta_E \log_\nu \frac{1}{\beta_E} \right] + o(1).
\]

At last, we prove 3). By employing the definition of \( M_E^{(0)} \) in (50), we can obtain

\[
\frac{\partial M_E^{(0)}}{\partial \alpha_E} = \frac{1}{1-2q} \ln \nu \left( 1 - \log_\nu \frac{1}{\beta_E} - \ln \frac{1}{\alpha_E} \right).
\]

Under Assumption 2, we know that \( \hat{q} > \frac{1}{2} \), and moreover, from (52), we know \( \nu > 1 \). Hence, \( \frac{\partial M_E^{(0)}}{\partial \alpha_E} < 0 \) if and only if

\[
1 - \log_\nu \frac{1}{\beta_E} - \ln \frac{1}{\alpha_E} < 0, \tag{64}
\]

which is equivalent to

\[
\alpha_E < \frac{1}{e} \left( \frac{1}{\beta_E} \right) \log_\nu \nu. \tag{65}
\]

Noting that \( \beta_E \in [0, 1] \) and \( \mu, \nu > 1 \), we know that if \( \alpha_E < \frac{1}{e} \), then (65) is satisfied, and hence, \( \frac{\partial M_E^{(0)}}{\partial \alpha_E} < 0 \).

On the other hand, from the definition of \( M_E^{(0)} \) in (50), we can obtain

\[
\frac{\partial M_E^{(0)}}{\partial \beta_E} = -\frac{1}{1-2q} \ln \mu \frac{1 - \alpha_E}{\beta_E} < 0. \tag{66}
\]

Similarly, it can be shown that for any given \( \psi_0 \) and \( \psi_1 \),

\[
\frac{\partial M_E^{(1)}}{\partial \alpha_E} < 0 \quad \text{and} \quad \frac{\partial M_E^{(1)}}{\partial \alpha_E} < 0.
\]

It is seen from the definitions in (50) and (51) that \( M_E^{(0)}, M_E^{(1)} \to \infty \) as \( \alpha_E, \beta_E \to 0 \), and hence, dominate the \( o(1) \) terms and determine the behavior of \( \mathbb{E}_0 \{ T_E \} \) and \( \mathbb{E}_1 \{ T_E \} \), respectively.

By comparing \( M_E^{(1)} \) in (50) and (51) with \( M_L^{(1)} \) in (13) and (14), it is clear that the dominant term \( M_E^{(1)} \) in \( \mathbb{E}_1 \{ T_E \} \) couples with the encryption parameters \( \psi_0 \) and \( \psi_1 \) in a very different way when compared to the dominant term \( M_L^{(1)} \) in \( \mathbb{E}_L \{ T_L \} \). In particular, \( M_L^{(1)} \) is determined by the KL divergences between the distributions under \( \mathcal{H}_0 \) and \( \mathcal{H}_1 \), while \( M_E^{(1)} \) is unrelated to the KL divergences.

Note that from Theorem 2, if \( \alpha_E < \frac{1}{e} \) and \( \beta_E < \frac{1}{e} \), then \( M_E^{(1)} \) monotonically decreases as \( \alpha_E \) or \( \beta_E \) increases. On the other hand, by employing the definitions of \( M_L^{(0)} \) and \( M_L^{(1)} \) in (13) and (14), after some algebra, we can obtain that if \( \alpha_L^* + \beta_L^* < 1 \), then

\[
\frac{\partial M_L^{(0)}}{\partial \alpha_L^*} \propto \frac{\partial M_L^{(1)}}{\partial \beta_L^*} \propto -\ln \left( 1 - \frac{\alpha_L^* - \beta_L^*}{\alpha_L^* \beta_L^*} \right) < 0, \tag{67}
\]

\[
\frac{\partial M_L^{(0)}}{\partial \beta_L^*} \propto \frac{\partial M_L^{(1)}}{\partial \alpha_L^*} \propto -\left( 1 - \alpha_L^* - \beta_L^* \right) < 0. \tag{68}
\]

which implies that for each \( i \), \( M_L^{(1)} \) also monotonically decreases as \( \alpha_L^* \) or \( \beta_L^* \) increases.

To compare the asymptotic performance of the SPRT with that of the MSPRT numerically, Fig. 1 depicts the values of \( M_L \triangleq \pi_0 M_L^{(0)} + \pi_1 M_L^{(1)} \) and \( M_E \triangleq \pi_0 M_E^{(0)} + \pi_1 M_E^{(1)} \) under different stochastic encryptions when \( \alpha_L^* = \beta_L^* = \alpha_E = \beta_E \) grow from \( 10^{-10} \) to \( 10^{-3} \). We set \( p = 0.7 \), \( q = 0.3 \), and the priors of \( \mathcal{H}_0 \) and \( \mathcal{H}_1 \) as \( \pi_0 = \pi_1 = 0.5 \). It is seen from Fig. 1 that \( M_L \) and \( M_E \) both decrease as \( \alpha_L^* = \beta_L^* = \alpha_E = \beta_E \) increase, and the difference between \( M_L \) and \( M_E \) vary considerably for different stochastic encryption parameters. For example, when \( \psi_0 = \psi_1 = 0.05 \), \( M_L = M_E \) which agrees with the results in Theorem 1, while there is a big gap between \( M_L \) and \( M_E \) when \( \psi_0 = 0 \) and \( \psi_1 = 0.2 \). Furthermore, Fig. 1 illustrates that under different stochastic encryptions, the difference between the slope of \( M_E \) and that of \( M_L \) can be significantly different. These observations motivate us to pursue the optimal stochastic encryption.

IV. OPTIMAL STOCHASTIC ENCRYPTION

In this section, we consider the optimization of the encryption parameters \( \psi_0 \) and \( \psi_1 \) with the goal of maximizing the difference between the expected sample sizes at the EFC and LFC with probabilities of false alarm and miss upper bounded by prescribed values. For a fair comparison between the expected sample sizes at the LFC and EFC, we set the upper bounds on the error probabilities to be identical at the LFC and EFC, i.e., \( \alpha_L < \alpha^* \), \( \alpha_E \leq \alpha^* \), \( \beta_L \leq \beta^* \), and \( \beta_E \leq \beta^* \). Moreover, denote \( \Psi \triangleq [\psi_0, \psi_1] \).

A. Optimization Formulation

To take into account the increase in the expected sample sizes at the LFC induced by the stochastic encryption, we impose the following constraints

\[
\lambda_i(\Psi) \triangleq \frac{\mathbb{E}_i \{ T_L \} - \mathbb{E}_i \{ T_E \}}{\mathbb{E}_i \{ T_L \}} \leq \kappa_i, i = 0, 1, \tag{69}
\]

where \( \kappa_i \) is a nonnegative constant which represents the upper bound on the acceptable tolerance of the increase in the expected sample sizes at the LFC induced by the stochastic encryption. The term \( \mathbb{E}_i \{ T_L \} \) in (69) corresponds to the case of no stochastic encryption, which can be obtained from (13) and (14) by replacing \( \bar{p} \) and \( \hat{q} \) by \( p \) and \( q \), respectively.

The motivation of the constraint in (69) deserves more discussion. In sequential detection, it is always desirable to reduce the expected sample size at the LFC as much as possible with the prescribed detection error performance guaranteed. If the employment of the stochastic encryption gives rise to a very large expected sample size at the LFC, the decision may not be meaningful in many practical scenarios where a prompt decision is critical, such as radar systems and smart grid systems. In light of this, the employment of the stochastic encryption should not increase the expected sample size at the LFC by a large amount.

Under Assumption 2, the optimization of the stochastic encryption parameter \( \Psi \) can be cast as the following maximin
problem

\[
\max_{\Psi, \alpha_L, \beta_L} \min_{\alpha_E, \beta_E} \sum_{i=0}^{1} \pi_i \left( \bar{E}_i \{ T_E \} - \bar{E}_i \{ T_L \} \right)
\]  \quad \text{s. t. } \lambda_i(\Psi) \leq \kappa_i, \forall i = 0, 1,
\]
\[
(1 - \psi_0 - \psi_1) p + \psi_0 > \frac{1}{2},
\]
\[
(1 - \psi_0 - \psi_1) q + \psi_0 < \frac{1}{2},
\]
\[
\alpha_L \leq \alpha^*, \beta_L \leq \beta^*,
\]
\[
\alpha_E \leq \alpha^*, \beta_E \leq \beta^*.
\]

Since the LFC employs the SPRT, it is clear that the constraint in (70e) is active for the optimal solution, that is, \( \alpha_L = \alpha^* \) and \( \beta_L = \beta^* \).

However, for the MSPRT, the expected sample sizes \( \bar{E}_i \{ T_E \} \) may not be minimized when \( \alpha_E = \alpha^* \) and \( \beta_E = \beta^* \) in general. To this end, as illustrated in the maximin problem in (70), we consider the best performance of the EFC in terms of the expected sample size when its detection performance satisfies the constraints, that is, \( \alpha_E \leq \alpha^* \) and \( \beta_E \leq \beta^* \).

B. Optimization Problem in Asymptotic Regime

In general, the optimization problem in (70) is not tractable, since the closed-form expression for the objective function in (70a) does not generally exist. In the following, we consider the objective function in (70a) in the asymptotic regime where \( \alpha^* \) and \( \beta^* \) are sufficiently small. From (13), (14), (50) and (51), by keeping the leading-order terms and ignoring the lower-order terms, \( \bar{E}_i \{ T_L \} \) and \( \bar{E}_i \{ T_E \} \) can be approximated in the asymptotic regime as

\[
\bar{E}_0 \{ T_L \} \approx -\ln \beta^* - \ln \frac{1}{\tilde{q}} + (1 - \tilde{q}) \ln \frac{1}{1 - \tilde{q}} \approx \tilde{t}_L^{(0)}(\Psi),
\]
\[
\bar{E}_1 \{ T_L \} \approx -\ln \alpha^* - \ln \frac{1}{\tilde{p}} + (1 - \tilde{p}) \ln \frac{1}{1 - \tilde{p}} \approx \tilde{t}_L^{(1)}(\Psi),
\]
\[
\bar{E}_0 \{ T_E \} \approx \frac{1}{1 - 2\tilde{q}} \log_\beta \frac{1}{\beta_E} \approx \tilde{t}_E^{(0)}(\Psi),
\]
and
\[
\bar{E}_1 \{ T_E \} \approx \frac{1}{2\tilde{p} - 1} \log_{\alpha_E} \frac{1}{\alpha_E} \approx \tilde{t}_E^{(1)}(\Psi).
\]

It is worth mentioning that the asymptotic approximations employed in (72)–(75) are similar to that suggested by Wald in [13] which is known as Wald’s approximation, and are commonly utilized in recent literature, see [33], [34] for instance.

It is seen from (74) and (75) that \( \tilde{t}_E^{(0)}(\Psi) \) and \( \tilde{t}_E^{(1)}(\Psi) \) are nonincreasing functions of \( \alpha_E \) and \( \beta_E \), and therefore, the optimization problem in (70) can be reduced to

\[
\max_{\Psi} \pi_0 \left[ \tilde{t}_E^{(0)}(\Psi) - \tilde{t}_L^{(0)}(\Psi) \right] + \pi_1 \left[ \tilde{t}_E^{(1)}(\Psi) - \tilde{t}_L^{(1)}(\Psi) \right]
\]  \quad (76a)
solution. In particular, both the objective function in (76a) and the feasible region specified by (69) is the corresponding asymptotic approximation of \( \lambda_i(\Psi) \) in (69).

By plugging (72)–(75) into (76), we attain an optimization problem where every term has an analytic expression. However, the optimization problem in (76) is generally nonconvex. In particular, the objective function in (76a) and the feasible region specified by (76b)–(76d) are generally nonconvex. Thus, it is generally intractable to find the globally optimal solution.

\[ \Psi(\lambda) \triangleq \{ \psi \in \mathbb{R}^n : \sum_{j=1}^n |\psi_j| \leq \lambda \}, \]

\( \lambda \) is the solution to (76) can be analytically given

\[ \lambda_i(\Psi) = \frac{\hat{T}_1(\Psi) - \hat{T}_1([0,0])}{\hat{T}_1([0,0])}, \quad i = 0, 1, \]

is the corresponding asymptotic approximation of \( \lambda_i(\Psi) \) in (69).

\section{C. Optimal Solution Under Small \( \kappa_i \)}

In this subsection, we will show that for small \( \kappa_i \) in (76b), the globally optimal solution to (76) can be analytically obtained.

We first look into the constraint in (76b). By employing (72) and (73), \( \bar{\lambda}_i(\Psi) \) can be rewritten as

\[ \lambda_0(\Psi) = \frac{H(q,p)}{H(p,p)} - 1 \quad \text{and} \quad \lambda_1(\Psi) = \frac{H(p,q)}{H(p,p)} - 1, \]

with \( H(x,y) \triangleq x \ln \frac{x}{y} + (1-x) \ln \frac{1-x}{1-y} > 0 \), if \( x \neq y \).

It is seen from (78) that the constraint in (76b) requires that \( H(\hat{p}, \hat{q}) \) and \( H(\hat{q}, \hat{p}) \) are nonzero, that is, \( \hat{p} \neq \hat{q} \).

The following Lemma provides some insights into the constraint in (76b). The proof is given in Appendix A.

\textbf{Lemma 1. Under Assumptions 1 and 2, we have the following results.}

1) For any given \( \Psi \),

\[ \nabla_{\Psi} \hat{\lambda}_i(\Psi) > 0, \quad i = 0, 1. \]

2) There exist two constants \( \zeta_\lambda \) and \( c_\lambda \) such that if

\[ \kappa \triangleq \max\{\kappa_0, \kappa_1\} < \zeta_\lambda \]

then

\[ \psi_j^{(\lambda)} \triangleq \sup \left\{ \psi_j \left| \begin{array}{l} \psi_j \| \psi_{-j} \|, s.t. \hat{\lambda}_i(\Psi) \leq \kappa_i, \quad i = 0, 1 \end{array} \right. \right\} \]

\[ = \sup \left\{ \psi_j \left| \hat{\lambda}_i(\Psi) \leq \kappa_i \quad \text{with} \quad \psi_{-j} = 0, \quad i = 0, 1 \right. \right\} \]

\[ = \min \left\{ \psi_j, \psi_{-j} \right\} \]

\[ < c_\lambda \kappa, \]

\begin{align*}
\text{where} & \quad \psi_{j,\hat{\lambda}_i} \text{ is the solution to } \hat{\lambda}_i(\Psi) = \kappa_i \text{ given } \psi_{-j} = 0. \\
\text{The constants } & \zeta_\lambda \text{ and } c_\lambda \text{ in Lemma 1 are defined in (129). Using the definition of } \psi_j^{(\lambda)} \text{ in (82), denote the following two points in the } \psi_0, \psi_1 \text{ plane,} \\
& \psi_0^{(\lambda)} = \left[ \psi_0^{(\lambda)}, 0 \right] \text{ and } \psi_1^{(\lambda)} = \left[ 0, \psi_1^{(\lambda)} \right]. \quad (86) \end{align*}

Noticing from (77), it is clear that \( \hat{\lambda}_0([0,0]) = 0, \ i = 0, 1. \) Moreover, as demonstrated by 1) in Lemma 1, \( \nabla_{\Psi} \hat{\lambda}_i(\Psi) > 0 \) which implies that for any \( \Psi_1 \) and \( \Psi_2 \), if \( \Psi_1 \succ \Psi_2 \), then \( \hat{\lambda}_i(\Psi_1) > \hat{\lambda}_i(\Psi_2) \), and hence \( \hat{T}_1(\Psi_1) > \hat{T}_1(\Psi_2) \), for all \( i = 0, 1. \) In other words, the larger \( \Psi \), the larger the expected sample size at the LFC. Thus, the stochastic encryption parameters should not be large in practice, since large stochastic encryption parameters give rise to a large expected detection delay at the LFC. It is worth mentioning that if \( \Psi \neq 0 \), then \( \hat{\lambda}_i(\Psi_1) > \hat{\lambda}_i([0,0]) = 0, \ i = 0, 1. \) and hence, \( \hat{T}_1(\Psi_1) > \hat{T}_1([0,0]), \ i = 0, 1. \) Therefore, every stochastic encryption degrades the performance of the SPRT at the LFC by increasing the expected sample size. Furthermore, as 2) in Lemma 1 illustrates, the points \( \Psi_0^{(\lambda)} \) and \( \Psi_1^{(\lambda)} \) respectively attain the largest possible values of \( \psi_0 \) and \( \psi_1 \) in the set specified by \( \cap_{i=0}^{n-1} \{ \hat{\lambda}_i(\Psi) \leq \kappa_i \}. \) Moreover, these two largest values \( \psi_0^{(\lambda)} \) and \( \psi_1^{(\lambda)} \) are bounded from above and can be controlled by \( \kappa \). It is worth mentioning that the set specified by \( \lambda_i(\Psi) \leq \kappa_i \) is just the region enclosed by \( [0, \psi_0^{(\lambda)}] \times \{0\}, \{0\} \times [0, \psi_1^{(\lambda)}] \) and the contour of \( \lambda_i(\Psi) = \kappa_i \).

Next we consider the constraints (76c) and (76d). Let \( \frac{\psi_0^{(\lambda)}}{1} \) and \( \frac{\psi_0^{(\lambda)}}{2} \) respectively denote the point where the line \( (1 - \psi_0 - \psi_1) p + \psi_0 = 1/2 \) intersects the \( \psi_1 \)-axis and the point where the line \( (1 - \psi_0 - \psi_1) q + \psi_0 = 1/2 \) intersects the \( \psi_0 \)-axis. It can be shown that

\[ \psi_1^{(\lambda)} = 1 - \frac{1}{2p} \quad \text{and} \quad \psi_0^{(\lambda)} = 1 - \frac{2q}{2(1-q)}. \]

According to the constraints in (76c) and (76d), we know that in the \( \psi_0, \psi_1 \) plane, the closed interval \( \{0\} \times [0, \psi_0^{(\lambda)}] \) and the closed interval \([0, \psi_0^{(\lambda)}] \times \{0\}\) are both contained in the set specified by (76c) and (76d). Therefore, if

\[ \kappa < \zeta_{\hat{p}, \hat{q}} \triangleq \min \left\{ \frac{\psi_0^{(\lambda)}}{\psi_1^{(\lambda)}}, \frac{\psi_1^{(\lambda)}}{\psi_0^{(\lambda)}} \right\}, \]

then \( \psi_0^{(\lambda)} \) and \( \psi_1^{(\lambda)} \) are bounded by \( \kappa \), and hence

\[ \sup \{ \psi_j \left| \begin{array}{l} \psi_j \| \psi_{-j} \|, s.t. \Psi = [\psi_{-j}, \psi_j] \in \mathcal{E} \end{array} \right. \right\} = \psi_j^{(\lambda)}, \]

where \( \mathcal{E} \) denotes the feasible set specified by all the constraints in the optimization problem in (76). From (89), we know that

\[ \mathcal{E} \subseteq \mathcal{E} \triangleq \left\{ \Psi : \psi_0 \in [0, \psi_0^{(\lambda)}], \psi_1 \in [0, \psi_1^{(\lambda)}] \right\}. \]

With regard to the behavior of the objective function in (76a), we have the following lemma. The proof is given in Appendix B.

\textbf{Lemma 2. Under Assumptions 1 and 2, we have the following results.}

\footnote{If } j = 0, \text{ then the vector } [\psi_{-j}, \psi_j] \text{ in (89) needs to be replaced by } [\psi_j, \psi_{-j}].
1) \( \hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi) \geq 0, \ i = 0, 1, \) with equality if and only if a symmetric encryption is employed, i.e., \( \psi_0 = \psi_1. \)

2) There exists a constant \( \zeta \) such that if \( \kappa < \zeta \),

\[
\kappa < \zeta_{\text{obj}},
\]

then in the region \( \mathcal{E} \cap \{ \Psi : \psi_1 > \psi_0 \geq 0 \} \), we have that

\[
\left\{ \begin{array}{ll}
\frac{\partial}{\partial \sigma_0} [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)] < 0, & \forall i = 0, 1,
\end{array} \right. \quad \text{(92)}
\]

while in the region \( \mathcal{E} \cap \{ \Psi : 0 < \psi_1 < \psi_0 \} \), we have that

\[
\left\{ \begin{array}{ll}
\frac{\partial}{\partial \sigma_0} [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)] > 0, & \forall i = 0, 1.
\end{array} \right. \quad \text{(93)}
\]

As illustrated by 1) in Lemma 2, the expected sample size at the EFC is no fewer than that at the LFC in the asymptotic regime where \( \alpha^*, \beta^* \to 0 \). From Theorem 1 and Lemma 2, we have the following corollary regarding the symmetric stochastic encryptions.

**Corollary 1.** Under Assumptions 1 and 2, the symmetric stochastic encryptions are the least favorable in the asymptotic regime where \( \alpha^*, \beta^* \to 0 \), since they are the only class of stochastic encryptions which cannot help the LFC outperform the EFC in terms of the expected sample sizes.

Finally, in the following theorem, for small \( \kappa \), we give the optimal stochastic encryption in the sense of maximizing the difference between the expected sample sizes at the EFC and LFC.

**Theorem 3.** If the following conditions

(C1) \( \psi_0^{(i)}(\lambda) \leq 1 - \frac{1-\pi}{2} \)

(C2) For all \( i = 0, 1, \)

\[
\frac{\partial}{\partial \sigma_0} [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)] < 0 \quad \text{and} \quad \frac{\partial}{\partial \sigma_1} [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)] > 0 \ \text{in the region} \ \mathcal{E} \cap \{ \Psi : \psi_1 > \psi_0 \geq 0, \}
\]

\[
\frac{\partial}{\partial \sigma_0} [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)] > 0 \quad \text{and} \quad \frac{\partial}{\partial \sigma_1} [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)] < 0 \ \text{in the region} \ \mathcal{E} \cap \{ \Psi : \psi_0 > \psi_1 \geq 0 \},
\]

hold, then

\[
\Psi^* = \arg \max_{\Psi \in \{ \Psi_0^{(i)}, \Psi_1^{(i)} \}} \sum_{i=0}^1 \pi_i \left[ \hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi) \right], \quad \text{(94)}
\]

where \( \Psi_0^{(i)} \) and \( \Psi_1^{(i)} \) are defined in (86).

Moreover, under Assumptions 1 and 2, the conditions in (C1) and (C2) hold provided that

\[
\kappa < \zeta^* \triangleq \min \left\{ \zeta_\lambda, \zeta_{\bar{\rho}, \bar{\eta}}, \zeta_{\text{obj}} \right\},
\]

where \( \zeta_\lambda \) and \( \zeta_{\text{obj}} \) are respectively defined in Lemma 1 and Lemma 2, and \( \zeta_{\bar{\rho}, \bar{\eta}} \) is defined in (88).

**Proof:** For the sake of notational simplicity, denote

\[
g_i(\psi_0, \psi_1) \triangleq \hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi) \quad \text{and} \quad g(\psi_0, \psi_1) \triangleq \sum_{i=0}^1 \pi_i [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)].
\]

If condition (C1) hold, then from (89), we have

\[
g_0(\psi_0^{(i)}(\lambda), \psi_1) = \sup \left\{ \psi_0 \mid \exists \psi_1, \text{ s.t.} \ \Psi = [\psi_0, \psi_1] \in \mathcal{E} \right\},
\]

and

\[
g_1(\psi_0^{(i)}(\lambda), \psi_1) = \sup \left\{ \psi_1 \mid \exists \psi_0, \text{ s.t.} \ \Psi = [\psi_0, \psi_1] \in \mathcal{E} \right\}.
\]

Hence, according to (C2), for any point \( \Psi = [\psi_0, \psi_1] \) in the region \( \mathcal{E} \cap \{ \Psi : \psi_1 > \psi_0 \geq 0 \}, \)

\[
g_i \left( \lambda, \psi_0^{(i)}(\lambda), \psi_1 \right) \geq g_i(0, \psi_1) \geq g_i(\psi_0, \psi_1),
\]

while in the region \( \mathcal{E} \cap \{ \Psi : \psi_0 > \psi_1 \geq 0 \}, \)

\[
g_i \left( \lambda, \psi_0^{(i)}(\lambda), 0 \right) \geq g_i(\psi_1, 0) \geq g_i(\psi_0, \psi_1).
\]

Therefore, in the region \( \mathcal{E} \cap \{ \Psi : \psi_1 > \psi_0 \geq 0 \}, \) we have

\[
g_0(0, \psi_0^{(i)}(\lambda)) = \sum_{i=0}^1 \pi_i g_i(0, \psi_0^{(i)}(\lambda)) = \max_{\Psi : \psi_1 > \psi_0} \sum_{i=0}^1 \pi_i \left( \hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi) \right) > 0,
\]

where the inequality in (100) is due to 1) in Lemma 2. Similarly, in the region \( \mathcal{E} \cap \{ \Psi : \psi_0 > \psi_1 \geq 0 \}, \) we have

\[
g_0(0, \psi_0^{(i)}(\lambda), 0) = \sum_{i=0}^1 \pi_i g_i(0, \psi_0^{(i)}(\lambda), 0) = \max_{\Psi : \psi_1 < \psi_0} \sum_{i=0}^1 \pi_i \left( \hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi) \right) > 0,
\]

where the inequality in (101) is also from 1) in Lemma 2. We conclude the proof for (94) by noting the fact that if \( \psi_1 = \psi_0 \), then

\[
\sum_{i=0}^1 \pi_i [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)] = 0 \quad \text{which is smaller than} \quad g_0(\psi_0^{(i)}(\lambda), 0) \quad \text{and} \quad g(0, \psi_0^{(i)}(\lambda)) \quad \text{from (100) and (101)}.
\]

Furthermore, by employing (90), Lemma 1 and Lemma 2, we know that if \( \kappa < \zeta^* \triangleq \min \left\{ \zeta_\lambda, \zeta_{\bar{\rho}, \bar{\eta}}, \zeta_{\text{obj}} \right\}, \) then the conditions (C1) and (C2) hold. ■

As illustrated in Theorem 3, there exists a constant \( \zeta^* \) such that if \( \kappa < \zeta^* \) then the optimal solution of the optimization problem in (76) can only be either the point \( \Psi_0^{(i)} \) or \( \Psi_1^{(i)} \). Thus, the optimization problem in (76) can be easily solved, though it is a nonconvex optimization problem. We summarize this procedure in Algorithm 1. The expression of \( \lambda_i(\Psi) \) can be found in (88). The expressions of

\[
\frac{\partial}{\partial \sigma_0} [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)] \quad \text{and} \quad \frac{\partial}{\partial \sigma_1} [\hat{T}_E^{(i)}(\Psi) - \hat{T}_L^{(i)}(\Psi)]
\]

can be found in (162) and (163), respectively. The closed-form expressions of the other partial derivatives in condition (C2) can be similarly obtained by following the steps for obtaining (162) and (163) in Appendix B. Thus, condition (C2) can be numerically evaluated. From (77), we can obtain \( \lambda_i([0, 0]) = 0, \ i = 0, 1. \) Moreover, by 1) in Lemma 1, we know that \( \lambda_i([0, 0]) \) are strictly increasing with respect to \( \psi_0 \) and \( \psi_1 \), respectively. Hence, if \( \kappa_i \) satisfies \( 0 \leq \kappa_i \leq \min \{ \lambda_i([0, 1]), \lambda_i([1, 0]) \} \), then by the Intermediate Value Theorem, we know that the solution \( \psi^{(i)}_0 \) in Step 3 in Algorithm 1 exists and is unique. In addition, by the monotonicity of \( \lambda_i([0, 0]) \) and \( \lambda_i([0, 1]) \),
\( \hat{\lambda}_i(\Psi) \) can be easily obtained by numerically searching the point along \( \psi_j \)-axis which achieves \( \hat{\lambda}_i(\Psi) = \kappa_i \) for \( i, j = 0, 1 \). It is worth mentioning that as demonstrated by Lemma 1 and Lemma 2, the conditions (C1) and (C2) can always be ensured provided that \( \kappa_0 \) and \( \kappa_1 \) are small enough. Note from (86), the optimal stochastic encryption strategy is just to flip one type of quantized bit with larger probability and keep the other type of quantized bit unchanged.

It is worth mentioning that the constant \( \zeta^* \) in (95) does not depend on \( \alpha^*, \beta^* \) and the stochastic encryption parameter \( \Psi \). Moreover, \( \Psi_{0}^{(\lambda)} \) and \( \Psi_{1}^{(\lambda)} \) do not depend on \( \alpha^*, \beta^*, \pi_0, \) and \( \pi_1 \). The optimal solution \( \Psi^* \) depends on \( \alpha^*, \beta^*, \pi_0, \) and \( \pi_1 \) only through the binary selection in (94). Here, we present an example to illustrate the results in Lemma 1, Lemma 2 and Theorem 3.

Consider the signal model in (5) where \( \theta = 1 \) and the independent noise \( W_k^{(n)} \sim N(0,1) \). The quantizers employed at the sensors are \( Q_n(x) = 1_{\{x \geq \frac{n}{2}\}} \) for all \( n \). In addition, the priors are \( \pi_0 = \pi_1 = 0.5 \) and the prescribed error probability bounds are \( \alpha^* = \beta^* = 10^{-6} \). Fig. 2 and Fig. 3 depict the contours of \( \hat{\lambda}_0(\Psi) \) and \( \hat{\lambda}_1(\Psi) \), respectively. It is seen that the feasible set specified by (76b)–(76d) in this case is nonconvex. Moreover, it is clear that \( \nabla \Psi \hat{\lambda}_i(\Psi) > 0 \) for \( i = 0, 1 \) which corroborates the results in Lemma 1. Fig. 4 illustrates the objective function \( \hat{T}_E(\Psi) - \hat{T}_L(\Psi) \triangleq \pi_0[\hat{T}_E^0(\Psi) - \hat{T}_L^0(\Psi)] + \pi_1[\hat{T}_E^1(\Psi) - \hat{T}_L^1(\Psi)] \) in (76a) versus the encryption parameters \( \psi_0 \) and \( \psi_1 \), and the contours of the objective function \( \hat{T}_E(\Psi) - \hat{T}_L(\Psi) \) is depicted in Fig. 5. As expected from Theorem 3, the numerical results in Fig. 4 verify that the maximum value of \( \hat{T}_E(\Psi) - \hat{T}_L(\Psi) \) can only be attained at either the upper left corner or the lower right corner. Furthermore, it is seen that the contour curves in Fig. 5 agree with the results in Lemma 2.

The implementation of the stochastic encryption in practice is worth some discussion. If we are only allowed to use \( N_{\text{bit}} \) digits to represent \( \psi_0 \) and \( \psi_1 \) in practice, then we need to round the optimal values of \( \psi_0 \) and \( \psi_1 \) to the nearest numbers which can be represented by \( N_{\text{bit}} \) digits, respectively. According to Lemma 1, if \( \psi_0 \) and \( \psi_1 \) are rounded toward positive infinity, then the constraint in (76b) can not be satisfied anymore. Thus, \( \psi_0 \) and \( \psi_1 \) can only be rounded toward negative infinity.
In light of this, as implied by Lemma 2 and Theorem 3, the optimal \( \psi_0 \) and \( \psi_1 \) under the constraint of an allowable number of digits can be obtained by truncating \( \psi_0 \) and \( \psi_1 \) after \( N_{\text{hit}} \) digits, respectively. The truncation can bring about some loss in the stochastic encryption performance which is defined by the objective function in (76a). The fewer the allowable number of digits, the larger the performance loss of the stochastic encryption. The required number of digits for representing \( \psi_0 \) and \( \psi_1 \) can be determined by the tolerance of the performance loss in practice.

\section*{D. Simulation Results}

In this subsection, we present a few simulation results to illustrate the performance of the optimal stochastic encryption.

The simulation setup considered in this subsection is the same as that for Fig. 2–Fig. 5. It is seen from Fig. 2–Fig. 5 that if \( 0 \leq \psi_0 \leq 0.2 \) and \( 0 \leq \psi_1 \leq 0.2 \), the conditions (C1) and (C2) hold. Assume that \( \kappa_0 = 0.265 \) and \( \kappa_1 = 0.2077 \). By employing Step 3 in Algorithm 1, we can obtain \( \Psi^{(\lambda)}_0 = [0.08, 0] \) and \( \Psi^{(\lambda)}_1 = [0, 0.1] \). Moreover, if \( \alpha^* = \beta^* \), then by employing (72)–(75), we can obtain \( \sum_{i=0}^{1} \pi_i [T^{(i)}_E((0, 0.1)] - T^{(i)}_L((0, 0.1)] = 1.756 \sum_{i=0}^{1} \pi_i [T^{(i)}_E((0, 0.08], [0.08, 0]) - T^{(i)}_L((0, 0.08], [0.08, 0])]. \) Hence, according to Theorem 3, \( \Psi^* = \Psi^{(\lambda)}_1 = [0, 0.1] \) is the optimal solution of (76) provided \( \alpha^* = \beta^* \). In the following, we compare the performance under different \( \Psi \), the optimal \( \Psi = [0, 0.1] \), some feasible but not optimal \( \Psi = [0, 0.05] \) and \( \Psi = [0.05, 0.05] \), and under no stochastic encryption, i.e., \( \Psi = [0, 0] \). The average sample sizes over \( 10^4 \) Monte Carlo runs at the LFC and EFC (i.e., \( \mathbb{E}\{T_L\} = \pi_0 \mathbb{E}_0\{T_L\} + \pi_1 \mathbb{E}_1\{T_L\} \) and \( \mathbb{E}\{T_E\} = \pi_0 \mathbb{E}_0\{T_E\} + \pi_1 \mathbb{E}_1\{T_E\} \) versus the prescribed error probability bounds (i.e., \( \alpha^* = \beta^* \)) are depicted in Fig. 6 and Fig. 7. As illustrated in Fig. 6, with no stochastic encryption, i.e., \( \Psi = [0, 0] \), then the expected sample sizes at the LFC and EFC are the same which agrees with the intuition. In addition, when the symmetric stochastic encryption \( \Psi = [0.05, 0.05] \) is employed, the simulation results in Fig. 6 verifies that the expected sample sizes at the LFC and EFC are the same as stated by Theorem 1. Moreover, the symmetric stochastic encryption \( \Psi = [0.05, 0.05] \) causes an increase in the expected sample size at the LFC compared to the case of no stochastic encryption, which verifies that the stochastic encryption incurs performance degradation at the LFC.

In Fig. 7, the performances of the optimal stochastic encryption \( \Psi = [0, 0.1] \) and the non-optimal stochastic encryption \( \Psi = [0, 0.05] \) are compared. It is seen that under the optimal stochastic encryption, the difference between the expected sample sizes at the EFC and LFC is significantly larger than that under the non-optimal one. However, this is at the price of an increase in the expected sample size at the LFC. In addition, the slope of \( \mathbb{E}\{T_E\} \) is smaller than that of \( \mathbb{E}\{T_L\} \), that is, as \( \alpha^* = \beta^* \) decreases, \( \mathbb{E}\{T_E\} \) grows faster than \( \mathbb{E}\{T_L\} \), and therefore, the difference between \( \mathbb{E}\{T_E\} \) and \( \mathbb{E}\{T_L\} \) becomes larger.

The assumption that the EFC employs the MSPRT deserves some discussion. The reasons why we make this assumption is mainly twofold. On one hand, if the EFC does not employ the MSPRT, then the LFC can easily defeat the EFC even without using any defending approach. This is because the LFC employs the SPRT which is the optimal sequential detection procedure, and hence even without employing any defending approach, the performance of the LFC is always better than that of the EFC provided that the EFC does not employ the SPRT in the absence of any defending approach. To this end,
the EFC has to employ the SPRT based on the distributions of the observations in the absence of any defending approach, so that the LFC needs to employ some defending approach, at the price of an increasing computational complexity or extra communication overhead or some other cost, to compete against the EFC. It is seen from (15) to (18) that the MSPRT employed at the EFC is exactly the SPRT based on the distributions of the observations in the absence of any defending approach. On the other hand, the EFC seems to have little option other than the MSPRT, since the distributions of the encrypted observations under two hypotheses are both unknown to the EFC. Moreover, it is not easy for the EFC to accurately estimate these two distributions. For example, one of the most viable approaches for the EFC to estimate the distribution of encrypted observations is to employ the maximum likelihood estimator, which can be expressed as

\[ \hat{D} \left( \left\{ \tilde{u}_t \right\}_{t=1}^{T_E} \right) \triangleq \frac{1}{T_E} \sum_{t=1}^{T_E} \tilde{u}_t, \]

where \( T_E \) is the number of observations received at the EFC. Due to the strong law of large numbers, we know that in the presence of the stochastic encryption, as \( T_E \rightarrow \infty \),

\[ \hat{D} \left( \left\{ \tilde{u}_t \right\}_{t=1}^{T_E} \right) \rightarrow \begin{cases} \tilde{p}, & \text{under } H_1, \\ \tilde{q}, & \text{under } H_0 \end{cases} \text{ almost surely.} \]

As implied by (103), in order to guarantee that the estimated distribution is close to the true distribution of the observations, the number \( T_E \) of observations must be sufficiently large. However, from the perspective of the EFC, it may not be worth waiting for a long time to accumulate enough observations to accurately estimate the distribution of encrypted observations, since the number of observations required for accurately estimating the distribution may be much larger than the expected number of observations required for the EFC to solve the sequential detection problem by employing the MSPRT. For example, as illustrated in Fig. 7, when \( \Psi = [0, 0.1] \) and the probabilities of false alarm and miss are required to be upper bounded by \( 10^{-5} \), which is a generally strict requirement on the sequential detection accuracy in practice, the expected number of observations required for the MSPRT to solve the sequential detection problem is only around 41, which is a pretty small number. On the other hand, for the case considered in Fig. 7, when \( \Psi = [0, 0.05] \), by employing (7) and (8), we can obtain \( p = 0.6915 \) and \( \tilde{p} = 0.6569 \), and hence \( |\tilde{p} - p| = 0.0346 \). However, as illustrated in Fig. 8, when \( H_1 \) is true and \( T_E \) is around 41, the expected difference between \( \hat{D}(\left\{ \tilde{u}_t \right\}_{t=1}^{T_E}) \) and \( \tilde{p} \) is more than 0.05 which is even larger than the difference between the distribution of encrypted data and that of unencrypted data. Thus, when compared with \( p \), the EFC cannot produce an estimated distribution with expectation closer to \( \tilde{p} \) by using such a small number of observations. Similarly, when \( H_0 \) is true, we also can show that when compared with \( q \), the EFC is unable to produce an estimated distribution with expectation closer to \( \tilde{q} \). In light of this, due to the poor estimation accuracy, it is not meaningful for the EFC to estimate the unknown distribution of the encrypted observations in such case.

Furthermore, since the EFC only receives data under one hypothesis, if the EFC wants to estimate the distribution of encrypted data under the other hypothesis, then the EFC needs to estimate the stochastic encryption parameters \( \psi_0 \) and \( \psi_1 \) and then employs (7) or (8) to produce an estimate. However, due to the fundamental limitation on the maximum parameter dimension for accurate estimation with quantized data [35], the EFC is unable to estimate these two stochastic encryption parameters \( \psi_0 \) and \( \psi_1 \) since it only receives i.i.d. one-bit data. Hence, the EFC cannot accurately estimate the distribution of encrypted data under any hypothesis. Therefore, it would be better for the EFC to just employ the MSPRT. It is worth pointing out that for the case where the EFC employs some other detection scheme, the stochastic encryption developed in this paper can still guarantee that the LFC is able to reach a decision before the EFC as long as they have the same detection accuracy.

In addition, the stochastic encryption and the EFC assumptions are both found in the related literature, see [16], [17], [21]–[23] for instance, and this work mainly focuses on the sequential detection aspect.
V. Conclusion

We have investigated sequential detection based on single-bit quantized data and in the presence of eavesdroppers. By employing stochastic encryptions at the sensors, each quantized bit is randomly flipped according to certain probabilities before transmitted to the LFC. The LFC knows both the distribution of the quantized data and the flipping probabilities and employs the optimal SPRT; whereas the EFC is unaware of the stochastic encryption and therefore employs a mismatched SPRT. We have characterized the expected sample size of the MSPT in terms of the detection thresholds. We have shown that when the detection error probabilities are set to be the same at the LFC and EFC, every symmetric stochastic encryption leads to the same expected size at the LFC and EFC. Furthermore, we have provided the asymptotic analysis on the expected sample size in terms of the vanishing error probabilities, and revealed the stark difference from the asymptotic performance of the SPRT with no model mismatch. In the asymptotic regime of small detection error probabilities, we have shown that every stochastic encryption degrades the performance of the SPRT at the LFC by increasing the expected sample size, and the expected sample size required at the EFC is no fewer than that required at the LFC. To this end, symmetric stochastic encryptions are the least favorable ones. Then we have considered the design of the optimal stochastic encryption in the sense of maximizing the difference between the expected sample size required at the EFC and LFC. Although this optimization problem is nonconvex, we have shown that if the acceptable tolerance of the increase in the expected sample size at the LFC induced by the stochastic encryption is small enough, the globally optimal stochastic encryption can be analytically obtained. Moreover, the optimal strategy randomly flips only one type of quantized bits (i.e., 0 or 1) and keeps the other type unchanged.

APPENDIX A

PROOF OF LEMMA 1

We first prove 1). From (78), we know that \( \nabla_{\Psi} \hat{\lambda}_i (\Psi) > 0, i = 0, 1 \) is equivalent to

\[
\nabla_{\Psi} H(\hat{p}, \hat{q}) < 0 \quad \text{and} \quad \nabla_{\Psi} H(\hat{q}, \hat{p}) < 0. \quad (104)
\]

By employing (79), after some algebra, we can obtain that

\[
\frac{\partial H(\hat{p}, \hat{q})}{\partial \psi_j} = \frac{\partial \hat{p}}{\partial \psi_j} \ln \frac{\hat{p} (1 - \hat{q})}{q (1 - \hat{p})} - \frac{\hat{p} - \hat{q}}{q (1 - \hat{q})} \frac{\partial \hat{q}}{\partial \psi_j}. \quad (105)
\]

Noticing \( \hat{p} > \hat{q} \) and plugging (7) and (8) into (105) yields

\[
\frac{\partial H(\hat{p}, \hat{q})}{\partial \psi_0} = (1 - p) \ln \frac{\hat{p} (1 - \hat{q})}{q (1 - \hat{p})} - (1 - q) \frac{\hat{p} - \hat{q}}{q (1 - \hat{q})} < 1 - p \left[ \frac{\hat{p} (1 - \hat{q})}{q (1 - \hat{p})} - 1 \right] - (1 - q) \frac{\hat{p} - \hat{q}}{q (1 - \hat{q})} \leq 0 \quad (106)
\]

\[
\frac{\partial H(\hat{p}, \hat{q})}{\partial \psi_1} = \frac{\hat{p} - \hat{q}}{\hat{q}} \left( 1 - \frac{p}{1 - p} \frac{1 - q}{1 - \hat{q}} \right), \quad (107)
\]

and

\[
\frac{\partial H(\hat{p}, \hat{q})}{\partial \psi_1} = -p \ln \frac{\hat{p} (1 - \hat{q})}{q (1 - \hat{p})} + q \frac{\hat{p} - \hat{q}}{q (1 - \hat{q})} \quad (108)
\]

\[
< -p \left[ 1 - \frac{\hat{q} (1 - \hat{p})}{\hat{p} (1 - \hat{q})} \right] + q \frac{\hat{p} - \hat{q}}{q (1 - \hat{q})} \quad (109)
\]

\[
= \frac{\hat{p} - \hat{q}}{1 - \hat{q}} \left( -\frac{\hat{p}}{\hat{p}} + \frac{\hat{q}}{\hat{q}} \right), \quad (110)
\]

where (106) and (109) are due to the fact that \( \ln x < x - 1 \) and \( \ln x > 1 - 1/x \) for all \( x > 1 \).

According to Assumption 1, we know that \( p > q \), and therefore, by employing (7) and (8), we can obtain

\[
\frac{p}{\hat{p}} = \frac{p}{1 - \psi_0 - \psi_1} \leq \frac{q (1 - \psi_0 - \psi_1) q + \psi_0}{1 - \psi_0 + \psi_1} = \frac{q}{\hat{q}},
\]

and

\[
\frac{1 - p}{1 - \hat{p}} = \frac{1}{1 - \psi_0 + \psi_1} \leq \frac{1}{1 - \psi_0 + \psi_1} = 1 - q, \quad (111)
\]

which yields

\[
\frac{\partial H(\hat{p}, \hat{q})}{\partial \psi_0} < 0 \quad \text{and} \quad \frac{\partial H(\hat{p}, \hat{q})}{\partial \psi_1} < 0,
\]

by employing (107) and (110). Similarly, we can prove \( \nabla_{\Psi} H(\hat{q}, \hat{p}) < 0 \) and hence, \( \nabla_{\Psi} \hat{\lambda}_i (\Psi) > 0, i = 0, 1 \).

Next, we will just prove 2) for \( \psi_1^{(\lambda)} \), and the proof for \( \psi_0^{(\lambda)} \) is similar.

It is clear that

\[
\psi_1^{(\lambda)} (\Psi) \leq \kappa_i \quad \text{with} \quad \psi_0 = 0, \quad i = 0, 1 \\quad \text{and} \quad \psi_1^{(\lambda)} (\Psi) \leq \kappa_i, \quad i = 0, 1 \quad \text{in the sense of maximizing the difference between the expected sample size required at the EFC and LFC. Although this optimization problem is nonconvex, we have shown that if the acceptable tolerance of the increase in the expected sample size at the LFC induced by the stochastic encryption is small enough, the globally optimal stochastic encryption can be analytically obtained. Moreover, the optimal strategy randomly flips only one type of quantized bits (i.e., 0 or 1) and keeps the other type unchanged.}
\]

\[
\psi_1^{(\lambda)} (\Psi) \leq \kappa_i \quad \text{with} \quad \psi_0 = 0, \quad i = 0, 1 \\quad \text{and} \quad \psi_1^{(\lambda)} (\Psi) \leq \kappa_i, \quad i = 0, 1 \quad \text{in the sense of maximizing the difference between the expected sample size required at the EFC and LFC. Although this optimization problem is nonconvex, we have shown that if the acceptable tolerance of the increase in the expected sample size at the LFC induced by the stochastic encryption is small enough, the globally optimal stochastic encryption can be analytically obtained. Moreover, the optimal strategy randomly flips only one type of quantized bits (i.e., 0 or 1) and keeps the other type unchanged.}
\]
If \( \psi_0 = 0 \) and \( \psi_1 < \frac{1}{2} \), then from (7), (8) and Assumption 1, we can obtain
\[
\frac{1}{4} \leq \frac{1}{2} \tilde{p} \leq p < 1, \quad (118)
\]
\[
0 < \frac{1}{2} q < \tilde{q} < q < \frac{1}{2}, \quad (119)
\]
and moreover,
\[
\tilde{p} = (1 - \psi_1) p > (1 - \psi_1) q = \tilde{q}, \quad (120)
\]
which imply that \( \frac{\partial H(q, \tilde{p})}{\partial \psi_1} \in (-\infty, 0) \) and \( H(q, \tilde{p}) \in (0, \infty) \). Furthermore, \( \frac{\partial \lambda_0(0, \psi_1)}{\partial \psi_1} \) and \( H(q, \tilde{p}) \) are continuous functions with respect to \( \psi_1 \) since \( \tilde{p} \) and \( \tilde{q} \) are continuous functions with respect to \( \psi_1 \). Therefore, if \( \psi_1 < \frac{1}{2} \), then
\[
\frac{\partial \lambda_0(0, \psi_1)}{\partial \psi_1} \text{ is a continuous function with respect to } \psi_1. \quad (124)
\]
Let
\[
d_0 = \left. \frac{\partial \lambda_0(0, \psi_1)}{\partial \psi_1} \right|_{\psi_1 = 0} > 0, \quad (121)
\]
where the inequality in (121) is due to the fact that \( \nabla \psi_1 \lambda_0(\Psi) > 0 \). Due to the continuity of \( \frac{\partial \lambda_0(0, \psi_1)}{\partial \psi_1} \) for \( \psi_1 < \frac{1}{2} \), there exists \( \zeta_0 \in (0, \frac{1}{2}) \) such that if \( \psi_1 < \zeta_0 \), then
\[
\frac{\partial \lambda_0(0, \psi_1)}{\partial \psi_1} \in \left( \left. \frac{d_0}{2}, \frac{3d_0}{4} \right) \right). \quad (122)
\]
Since \( \lambda_0([0, 0]) = 0 \) and \( \lambda_0([0, \psi_1]) \) is an increasing function with respect to \( \psi_1 \), we know that \( \lambda_0([0, \zeta_0]) > \frac{d_0}{2} \zeta_0 \) due to (122). Hence, if we choose \( \kappa_0 < \frac{d_0}{2} \zeta_0 \), then \( \psi_1, \lambda_0 < \zeta_0 \). Moreover, if \( \psi_1, \lambda_0 < \zeta_0 \), then from (122), we know that
\[
\frac{d_0}{2} \psi_1, \lambda_0 < \lambda_0([0, \psi_1, \lambda_0]) \leq \kappa_0, \quad (123)
\]
which implies that
\[
\psi_1, \lambda_0 < \frac{2d_0}{d_0} \kappa_0 = \frac{2d_0}{d_0} \kappa, \quad (124)
\]
Similarly, there exist two constants \( d_1 \) and \( \zeta_1 \) such that if we choose \( \kappa_1 < \frac{d_1}{2} \zeta_1 \), then \( \psi_1, \lambda_0 < \frac{d_1}{2} \kappa_1 \), which implies that if \( \kappa < c_1^{(1)}_\lambda \), then
\[
\psi_1^{(1)} = \min \left\{ \psi_1, \lambda_0, \psi_1, \lambda_1 \right\} < c_1^{(1)}_\lambda \kappa, \quad (125)
\]
where \( c_1^{(1)}_\lambda \) and \( c_1^{(1)} \) are defined as
\[
c_1^{(1)}_\lambda \triangleq \min \left\{ \frac{d_0}{2} \zeta_0, \frac{d_1}{2} \zeta_1 \right\}, \quad (126)
\]
\[
c_1^{(1)} \triangleq \max \left\{ \frac{2d_0}{d_0} \frac{2d_0}{d_1} \right\}. \quad (127)
\]
Analogous to (126) and (127), there exist two constants \( c_0^{(0)} \) and \( c_0^{(0)}_\lambda \), such that if \( \kappa < c_0^{(0)}_\lambda \), then
\[
\psi_0^{(0)} = \min \left\{ \psi_0, \lambda_0, \psi_0, \lambda_1 \right\} < c_0^{(0)} \kappa. \quad (128)
\]
Therefore, by defining
\[
\zeta_\lambda \triangleq \min \left\{ \frac{c_0^{(0)}_\lambda}{c_0^{(0)}}, \frac{c_1^{(1)}_\lambda}{c_1^{(1)}} \right\} \quad \text{and} \quad \zeta_\lambda \triangleq \max \left\{ \frac{c_0^{(0)}}{c_0^{(0)}}, \frac{c_1^{(1)}}{c_1^{(1)}} \right\}, \quad (129)
\]
we obtain (85) from (125) and (128).

\footnote{In what follows, \( \frac{\partial \lambda_0(0, \psi_1)}{\partial \psi_1} \triangleq \frac{\partial \lambda_0(0, \psi_0, \psi_1)}{\partial \psi_1} \right|_{\psi_0 = 0}. \}

**APPENDIX B**

**PROOF OF LEMMA 2**

We first prove 1). By employing (72) and (74), we can obtain
\[
\hat{T}_E^{(0)}(\Psi) - \hat{T}_L^{(0)}(\Psi) = f(\Psi) \ln \frac{1}{\beta^r}, \quad (130)
\]
with \( f(\Psi) \triangleq \left\{ \left(1 - 2\tilde{q}\right) \left[ \ln \frac{\tilde{p}}{\ln (1 - \tilde{p})} \right]^{-1} - \frac{1}{H(\tilde{q}, \tilde{p})} \right\} \triangleq \psi_{G(\tilde{q}, \tilde{p})}, \quad (131)
\]
Notice that
\[
G(\tilde{q}, \tilde{p}) - H(\tilde{q}, \tilde{p}) = (1 - 2\tilde{q}) \ln \tilde{p} - \ln \tilde{q} \ln \frac{\tilde{q}}{\tilde{p}} + (1 - \tilde{q}) \ln 1 - \tilde{q}, \quad (132)
\]
where the inside of the bracket in (132) is the KL divergence of two Bernoulli distributions, and therefore,
\[
G(\tilde{q}, \tilde{p}) \leq H(\tilde{q}, \tilde{p}), \quad (133)
\]
with equality if and only if \( \tilde{p} + \tilde{q} = 1 \). Hence, from (7) and (8), we know that
\[
\hat{T}_E^{(0)}(\Psi) - \hat{T}_L^{(0)}(\Psi) \geq 0, \quad (134)
\]
with equality if and only if \( \psi_0 = \psi_1 \), which proves 1) for \( T_E^{(1)}(\Psi) - \hat{T}_L^{(1)}(\Psi) \). The proof for \( T_E^{(1)}(\Psi) - \hat{T}_L^{(1)}(\Psi) \geq 0 \) with equality if and only if \( \psi_0 = \psi_1 \) is similar.

Next, we consider 2). We first prove (92) for \( \hat{T}_E^{(0)}(\Psi) - \hat{T}_L^{(0)}(\Psi) \) in the region \( \Gamma \times \{ \psi_1 > \psi_0 \geq 0 \} \). Denote \( \delta \triangleq \psi_1 - \psi_0 > 0 \). Under Assumption 1, by employing (7) and (8), we can obtain
\[
\tilde{p} = 1 - \tilde{q} - \delta, \quad (135)
\]
By taking partial derivative of \( f(\Psi) \) with respect to \( \psi_j \), after some algebra, we can obtain
\[
\frac{\partial f(\Psi)}{\partial \psi_j} = \left[ \frac{1}{H(\tilde{q}, \tilde{p})^2} \frac{\tilde{p} (1 - \tilde{p})}{(1 - \tilde{p})} - \frac{1}{G(\tilde{q}, \tilde{p})^2} \frac{1 - 2\tilde{q}}{\tilde{p} (1 - \tilde{p})} \right] \frac{\partial \tilde{p}}{\partial \psi_j} \triangleq \psi_{Y_1(\tilde{q}, \delta)}, \quad (136)
\]
and
\[
\frac{\partial f(\Psi)}{\partial \psi_j} = \left[ \frac{2}{G(\tilde{q}, \tilde{p})^2} \ln \tilde{q} - \frac{1}{H(\tilde{q}, \tilde{p})} \ln \frac{\tilde{p} (1 - \tilde{q})}{\tilde{q} (1 - \tilde{p})} \right] \frac{\partial \tilde{q}}{\partial \psi_j} \triangleq \psi_{Y_2(\tilde{q}, \delta)}, \quad (137)
\]
From (135), we know that \( \tilde{p} < 1 - \tilde{q} \) and \( 1 - 2\tilde{q} > \tilde{p} - \tilde{q} \), and hence
\[
Y_1(\tilde{q}, \delta) < 0, \quad (137)
\]
since \( G(\tilde{q}, \tilde{p}) < H(\tilde{q}, \tilde{p}) \) as illustrated in (133).
On the other hand, $Y_2(\hat{q}, \delta)$ can be rewritten as

$$Y_2(\hat{q}, \delta) = \frac{1}{(1 - 2\hat{q})^2} H(\hat{q}, \hat{p})^2 \ln \frac{\hat{p}}{\hat{q}} \Bigg|_{\hat{Z}_2(\hat{q}, \delta)}$$

$$\times \left[ 2H(\hat{q}, \hat{p})^2 \ln \frac{\hat{p}(1 - \hat{q})}{\hat{q}(1 - \hat{p})} \Bigg|_{\hat{Z}_2(\hat{q}, \delta)} \right]$$

which implies

$$Y_2(\hat{q}, 0) = Z_1(\hat{q}, 0) - Z_2(\hat{q}, 0) = 0.$$  \hspace{1cm} (139)

Furthermore, by taking partial derivative of $Z_1(\hat{q}, \delta)$ and $Z_2(\hat{q}, \delta)$ with respect to $\delta$, we can obtain

$$\frac{\partial Z_1(\hat{q}, \delta)}{\partial \delta} = 2H(\hat{q}, 1 - \hat{q} - \delta) \frac{\partial H(\hat{q}, 1 - \hat{q} - \delta)}{\partial \delta} = \frac{4H(\hat{q}, 1 - \hat{q} - \delta)}{(1 - \hat{q} - \delta)(\hat{q} + \delta)} (2\hat{q} + \delta - 1) < 0,$$

since $2\hat{q} + \delta - 1 = \hat{q} - \hat{p} < 0$ according to Assumption 2, and

$$\frac{\partial Z_2(\hat{q}, \delta)}{\partial \delta} = \left[ \ln \frac{\hat{p}(1 - \hat{q})}{\hat{q}(1 - \hat{p})} \frac{\partial \ln \frac{\hat{p}}{\hat{q}}}{\partial \delta} \left|_{\hat{Z}_2(\hat{q}, \delta)} \right] + \ln \frac{\hat{p}(1 - \hat{q})}{\hat{q}(1 - \hat{p})} \frac{\partial \ln \frac{\hat{p}(1 - \hat{q})}{\hat{q}(1 - \hat{p})}}{\partial \delta} \right]$$

which yields

$$d_Z(\hat{q}, \delta) = \frac{\partial [Z_1(\hat{q}, \delta) - Z_2(\hat{q}, \delta)]}{\partial \delta} \frac{Z_1(\hat{q}, \delta)}{(1 - \hat{q} - \delta)(\hat{q} + \delta)},$$

with $Z_1(\hat{q}, \delta) = -4(1 - 2\hat{q} - \delta) \frac{(1 - \hat{q} - \delta)H(\hat{q}, 1 - \hat{q} - \delta)}{(1 - \hat{q} - \delta)(\hat{q} + \delta)} + (1 - 2\hat{q})^2 \left( 2\ln \frac{1 - \hat{q} - \delta}{\hat{q} + \delta} + \ln \frac{1 - \hat{q}}{\hat{q}} \right)$.

Note that

$$Z_1(\hat{q}, 0) = -(1 - 2\hat{q})^2 \ln \frac{1 - \hat{q}}{\hat{q}} < 0,$$  \hspace{1cm} (144)

since $\hat{q} < 0.5$ according to Assumption 2. Hence, from (142), we can obtain that for all $\hat{q} \in (0, \frac{1}{2})$,

$$d_Z(\hat{q}, 0) = -(1 - 2\hat{q})^2 \ln \frac{1 - \hat{q}}{\hat{q}} < 0$$  \hspace{1cm} (145)

since $1 - \hat{q} > 0$ and $\hat{q} > 0$.

Define a constant $\omega_0$

$$\omega_0 = \frac{1 - 2\hat{q}}{4c_3(1 - \hat{q})},$$

where $c_3$ is the constant in (85).

By employing Lemma 1, we know that if

$$\kappa < \min \{\zeta_\lambda, \omega_0\},$$

then $\psi_0 < c_\lambda \kappa$ and $\psi_1 < c_\lambda \kappa$, and therefore, from (8), we can obtain

$$\hat{q} = q + (1 - q)\psi_0 - \psi_1q \in [q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)]$$

$$\subset \left[ q - \frac{1 - 2q}{4(1 - q)}q, \frac{1}{4} + \frac{1}{2}q \right]$$

$$\subset \left[ 0, \frac{1}{2} \right],$$

where (149) and (150) are due to (146), (147) and the fact that $0 < q < 0.5$. From (145), we know that $d_Z(\hat{q}, 0)$ is a continuous function with respect to $\hat{q}$ over $\hat{q} \in (0, \frac{1}{2})$, and hence, can achieve its maximum $d_Z^* < 0$ over the closed set $[q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)]$, that is,

$$d_Z(\hat{q}, 0) \leq d_Z^* < 0, \forall \hat{q} \in [q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)].$$  \hspace{1cm} (151)

Notice from (142) that $d_Z(\hat{q}, \delta)$ is continuous with respect to $\hat{q}$, $\delta$, and $\kappa$. Thus, from (145), we know that $\forall \hat{q} \in [q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)]$, there exists $\varepsilon_\hat{q} > 0$ such that if

$$(\hat{q}, \delta) \in (\hat{q} - \varepsilon_\hat{q}, \hat{q} + \varepsilon_\hat{q}) \times [0, \varepsilon_\hat{q})$$

then

$$d_Z(\hat{q}, \delta) < \frac{3}{2} d_Z(\hat{q}, 0), \frac{1}{2} d_Z(\hat{q}, 0).$$

Noting that

$$[q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)] \subset \bigcup \left( \hat{q} \in [q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)] \right),$$

and the set $[q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)]$ is compact, we know that there exist $\{\hat{q}_1, \hat{q}_2, ..., \hat{q}_M\} \subset [q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)]$ such that

$$[q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)] \subset \bigcup_{i=1}^M (\hat{q}_i, \hat{q}_i + \varepsilon_\hat{q}_i).$$

By defining $\varepsilon_\hat{q} \triangleq \min \{\varepsilon_\hat{q}_1, \varepsilon_\hat{q}_2, ..., \varepsilon_\hat{q}_M\}$, we can obtain from (153) that if $\delta \in [0, \varepsilon_\hat{q}]$, then $\forall \hat{q} \in [q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)]$,

$$d_Z(\hat{q}, \delta) < \frac{1}{2} d_Z(\hat{q}, 0) \leq \frac{1}{2} d_Z^* < 0.$$

Let $\omega_1$ denote a constant

$$\omega_1 \triangleq \frac{\varepsilon_\hat{q}}{c_\lambda}.$$  \hspace{1cm} (156)

If $\kappa < \min \{\zeta_\lambda, \omega_0, \omega_1\}$, then by employing Lemma 1, we can obtain that $\forall \hat{q} \in [\hat{q}_1, \hat{q}_2]$,

$$\psi_0 \in [0, \varepsilon_\hat{q}]$$

and hence, $\delta = \psi_1 - \psi_0 \in (0, \varepsilon_\hat{q})$, which implies $\forall \hat{q} \in [q - c_\lambda \kappa q, q + c_\lambda \kappa (1 - q)]$ and $\forall \delta \in (0, \varepsilon_\hat{q})$, by employing Taylor’s theorem, there exists a $\delta \in (0, \delta)$ such that

$$Z_1(\hat{q}, \delta) - Z_2(\hat{q}, \delta) = [Z_1(\hat{q}, 0) - Z_2(\hat{q}, 0)] + d_Z(\hat{q}, \delta)\delta$$

$$\leq \frac{1}{2} d_Z^* \delta < 0$$

(158)
from (139) and (155). As a result, by employing (138), we can obtain
\[ Y_2(\tilde{\eta}, \tilde{\delta}) = Z_0(\tilde{\eta}, \tilde{\delta})(Z_1(\tilde{\eta}, \tilde{\delta}) - Z_2(\tilde{\eta}, \tilde{\delta})) < 0, \]  
(159)
since \( Z_0(\tilde{\eta}, \tilde{\delta}) > 0 \) which is a consequence of \( \tilde{p} > 0.5 > \tilde{q} \) according to Assumption 2.

Furthermore, from (7) and (8), we can obtain
\[ \frac{\partial \tilde{L}}{\partial \tilde{q}} = - \frac{\partial q}{\psi_0} = 1 - p \quad \text{and} \quad \frac{\partial \tilde{L}}{\partial \psi_0} = - \frac{\partial \tilde{L}}{\partial \psi_1} = p. \]  
(160)
Therefore, by defining
\[ \rho_1 \triangleq \min \{ \zeta_{\lambda}, \tilde{\omega}_0, \omega_1 \}, \]  
(161)
and by employing (130), (136), (137), (159) and (160), we know that if \( \kappa < \rho_1 \), then in the region \( \tilde{E} \cap \{ \Psi : \psi_1 > \psi_0 \geq 0 \} \),
\[ \frac{\partial}{\partial \psi_0} \tilde{L}^{(0)}(\Psi) - \tilde{L}^{(0)}(\Psi) \]
\[ = \frac{1}{\beta^*}, \]
(162)
and
\[ \frac{\partial}{\partial \psi_1} \tilde{L}^{(0)}(\Psi) - \tilde{L}^{(0)}(\Psi) \]
\[ = \frac{1}{\beta^*}, \]
(163)
since \( \beta^* < 1 \), which complete the proof of (92) for \( \tilde{E}^{(0)}(\Psi) - \tilde{L}^{(0)}(\Psi) \).

By following a similar approach as above, we can show that in the region \( \tilde{E} \cap \{ \Psi : 0 \leq \psi_1 < \psi_0 \} \), there exists a constant \( \rho_0 \) such that if \( \kappa < \rho_0 \), then (93) for \( \tilde{E}^{(0)}(\Psi) - \tilde{L}^{(0)}(\Psi) \) is true. Let \( \zeta_{\rho_0}^{(0)} \triangleq \min \{ \rho_0, \rho_1 \} \).

Furthermore, we can similarly prove that there exists a constant \( \zeta_{\rho_1}^{(0)} \) such that if \( \kappa < \zeta_{\rho_1}^{(0)} \), then the corresponding results in (92) and (93) for \( \tilde{E}^{(0)}(\Psi) - \tilde{L}^{(0)}(\Psi) \) are true. For the sake of brevity, the details of these proofs are omitted. Finally, we can conclude the proof for 2) by defining
\[ \zeta_{\rho_1}^{(0)} \triangleq \min \{ \zeta_{\rho_1}^{(0)}, \zeta_{\rho_1}^{(1)} \}. \]  
(164)
[33] K. Cohen and Q. Zhao, “Asymptotically optimal anomaly detection via sequential testing,” *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2929–2941, Jun. 2015.

[34] C.-Z. Bai, V. Katewa, V. Gupta, and Y.-F. Huang, “A stochastic sensor selection scheme for sequential hypothesis testing with multiple sensors,” *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3687–3699, Jul. 2015.

[35] J. Zhang, R. S. Blum, L. M. Kaplan, and X. Lu, “A fundamental limitation on maximum parameter dimension for accurate estimation with quantized data,” *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6180–6195, Sep. 2018.

**Jiangfan Zhang** (S’10–M’17) received the B.Eng. degree in communication engineering from Huazhong University of Science and Technology, Wuhan, China, in 2008, the M.Eng. degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2011, and the Ph.D. degree in electrical engineering from Lehigh University, Bethlehem, PA, USA, in 2016.

From 2016 to 2018, he was a Post-Doctoral Research Scientist with the Department of Electrical Engineering, Columbia University, New York, NY, USA. Since 2018, he has been with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO, USA, where he is currently an Assistant Professor. His research interests include signal processing, machine learning, and their applications to cybersecurity, cyber-physical systems, Internet of Things, sensor networks, smart grid, and sonar processing. He was a recipient of the Dean’s Doctoral Student Assistantship, the Gotshall Fellowship and the P. C. Rossin Doctoral Fellowship at Lehigh University.

**Xiaodong Wang** (S’98–M’98–SM’04–F’08) received the Ph.D. degree in electrical engineering from Princeton University.

He is currently a Professor of electrical engineering at the Columbia University in the City of New York. He has published extensively in these areas. Among his publications is a book titled *Wireless Communication Systems: Advanced Techniques for Signal Reception*, (Prentice Hall, 2003). His research interests fall in the general areas of computing, signal processing, and communications. He is listed as an ISI Highly-cited Author. His current research interests include wireless communications, statistical signal processing, and genomic signal processing. He received the 1999 NSF CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He has served as an Associate Editor for the *IEEE Transactions on Communications*, the *IEEE Transactions on Wireless Communications*, the *IEEE Transactions on Signal Processing*, and the *IEEE Transactions on Information Theory*. 