

Article

# Fundamental Limitations in Energy Detection for Spectrum Sensing

Xiao-Li Hu <sup>1,2</sup> , Pin-Han Ho <sup>3</sup> and Limei Peng <sup>4,\*</sup> <sup>1</sup> Nanfang College, Sun Yat-Sen University, Guangzhou 510900, China; hutby@yeah.net<sup>2</sup> College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China<sup>3</sup> Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada; p4ho@uwaterloo.ca<sup>4</sup> School of Computer Science and Engineering, Kyungpook National University, Daegu 41566, Korea

\* Correspondence: aurorapl@knu.ac.kr; Tel.: +82-053-950-5553

Received: 15 May 2018; Accepted: 27 June 2018; Published: 28 June 2018



**Abstract:** A key enabler for Cognitive Radio (CR) is spectrum sensing, which is physically implemented by sensor and actuator networks typically using the popular energy detection method. The threshold of the binary hypothesis for energy detection is generally determined by using the principles of constant false alarm rate (CFAR) or constant detection rate (CDR). The CDR principle guarantees the CR primary users at a designated low level of interferences, which is nonetheless subject to low spectrum usability of secondary users in a given sensing latency. On the other hand, the CFAR principle ensures secondary users' spectrum utilization at a designated high level, while may nonetheless lead to a high level of interference to the primary users. The paper introduces a novel framework of energy detection for CR spectrum sensing, aiming to initiate a graceful compromise between the two reported principles. The proposed framework takes advantage of the summation of the false alarm probability  $P_{fa}$  from CFAR and the missed detection probability  $(1 - P_d)$  from CDR, which is further compared with a predetermined confidence level. Optimization presentations for the proposed framework to determine some key parameters are developed and analyzed. We identify two fundamental limitations that appear in spectrum sensing, which further define the relationship among the sample data size for detection, detection time, and signal-to-noise ratio (SNR). We claim that the proposed framework of energy detection yields merits in practical policymaking for detection time and design sample rate on specific channels to achieve better efficiency and less interferences.

**Keywords:** fundamental limitations; SNR; energy detection; noise variance; spectrum sensing

## 1. Introduction

Cognitive Radio (CR) was first introduced as a promising candidate for dealing with spectrum scarcity in future wireless communications [1]. Under-utilized frequency bands originally allocated to licensed users (i.e., primary users) are freed and become accessible by non-licensed users (i.e., secondary users) equipped with CR in an opportunistic manner to maximize the spectrum utilization while minimizing interferences to the primary users. Despite its obvious advantages, CR technology is subject to a great challenge in detection of spectrum holes through spectrum sensing. This is because secondary users generally have very limited knowledge about the whole spectrum, which may leave the spectrum sensing results far from accurate. Some existing spectrum sensing methods in the literature are by way of matched filtering, waveform-based sensing [2], cyclostationary-based sensing [3,4], eigenvalue-based method [5,6], energy detection [7–13], etc. Obviously, energy detection is the most popular and simple way for spectrum sensing.

The energy detection method [7–13] for spectrum sensing measures the average energy of the total received signal during a period of time and compares it with a properly assigned threshold to decide the presence or absence of users. Typically, the energy detection method is formulated in a binary hypothesis test with a null Hypothesis  $H_0$  for absence of users and an alternative Hypothesis  $H_1$  for presence. The threshold is determined typically based on two standard principles: constant false alarm rate (CFAR) and constant detection rate (CDR) [14,15]. With the emphasis on promoting usage of spectrum hole, a threshold by CFAR is derived by assuring the probability of false alarm under  $H_0$  less than a given confidence level  $\alpha$ , while, with the emphasis on less interference of users, a threshold by CDR is derived by letting the probability of missed detection under  $H_1$  less than a given confidence level  $\alpha$ . Each criterion can ensure the error detection probability of one hypothesis under a low level while ignoring the error detection probability for the opponent hypothesis. Therefore, under some extreme circumstances, one error probability may be large although the other one is small. It is the purpose of this paper to develop a new criterion to simultaneously keep the two kinds of error detection probabilities at a low level. To the author’s knowledge, it is the first effort to develop this kind of criterion in energy detection for spectrum sensing.

A simple way to keep the two kinds of error detection probabilities simultaneously small is to restrict the summation of the two probabilities to less than a confidence level. To describe more in detail, let  $u$  be a constructed statistics for the binary hypothesis test. For a given small  $\alpha \in (0, 1)$ , say  $\alpha = 0.05$ , the threshold by CFAR principle is selected by the smallest  $\lambda$  such that

$$P[u \geq \lambda|H_0] \leq \alpha,$$

while the threshold by CDR principle is selected by maximal  $\lambda$  such that

$$P[u \leq \lambda|H_1] \leq \alpha.$$

The new criterion to select  $\lambda$  based on the summation of the two probabilities is proposed as

$$P[u \geq \lambda|H_0] + P[u \leq \lambda|H_1] \leq \alpha,$$

which ensures the two probabilities simultaneously smaller than the confidence level  $\alpha$ . Denote the false alarm probability  $P_{fa}(\lambda) = P[u \geq \lambda|H_0]$  for CFAR and the missed detection probability  $(1 - P_d(\lambda))$  for CDR with  $P_d(\lambda) = P[u \leq \lambda|H_1]$ , respectively. The aforementioned summation principle turns to be

$$P_{fa}(\lambda) + (1 - P_d(\lambda)) \leq \alpha, \tag{1}$$

This is actually not a well-posed presentation though, since there may be too many solutions or no solution sometimes to the inequality in Equation (1) with respect to  $\lambda$  for a given noise variance  $\sigma_n^2$ , signal to noise ratio (SNR), and data size  $M$ .

In this paper, to derive unique solution of threshold  $\lambda$  from (1), two kinds of optimization problems are introduced: (i) to find the minimum data size  $M$  with  $\sigma_n^2$  and SNR given; and (ii) to find the lowest SNR with  $\sigma_n^2$  and  $M$  given. Under the first optimization setting, we show that, with a small SNR, the data size  $M$  should be larger than a critical value, denoted by  $M_{\min}$ , to guarantee the existence of threshold  $\lambda$  that can satisfy the inequality in Equation (1) under a given confidence level  $\alpha$ . We also show an asymptotical formula, i.e., Equation (61), for the minimum data size  $M_{\min}$  as SNR goes to zero. Under the second optimization setting, we show that for a given data size  $M$  the SNR should be greater than a minimum SNR to ensure the existence of threshold  $\lambda$  that satisfies inequality in Equation (1) under given confidence level  $\alpha$ . The asymptotical formula for the minimum SNR, i.e., Equation (62), is found as  $M \rightarrow \infty$ . Theoretical analysis and simulations are conducted for the two optimization settings and the obtained asymptotical formulas are verified.

The main contributions of this paper are as follows. (i) A new principle to select the threshold of energy detection is proposed by assuring the summation of the two kind of error detection probabilities

less than a given confidence level. To derive unique threshold under the constraint, two kinds of optimization frameworks are proposed. (ii) The possible optimal selection of the thresholds under the proposed two optimization frameworks have been analyzed in Propositions 1 and 3 regardless of the constraints. (iii) The lower and upper bounds of the solutions for the two optimization frameworks have been established in Theorems 1–4, respectively, to the accurate distribution and the approximate distribution of the test statistics. Two asymptotical formulae under corresponding limit process for the two optimization problems are also derived to describe the fundamental limitations when using energy detection for spectrum sensing.

The fundamental limitations in energy detection found in this paper based on the constraint Equation (1) are different from the SNR wall introduced in [16], which is on the other hand a limitation regarding robustness of detectors. It is discovered in this paper that even when the noise variance keeps constant, some limitations still exist regarding the tradeoff between efficiency and noninterference. For example, when channel detection time is 2 s [17] and sample rate of a channel [18] is once every 16 μs (which yields the data size  $M = 2/0.000016 = 125,000$ ), by the asymptotical Equation (62), the minimum SNR is approximately 0.0158 (i.e., −18.0124 dB) under a confidence level  $\alpha = 0.05$ . In other words, it is impossible to detect any signal with SNR lower than −18.0124 dB under such a setting with a confidence level  $\alpha = 0.05$ . The analysis and understanding on these limitations not only enables a wise choice of channels in the CR spectrum sensing, but also helps policymaking in determination of detection settings, such as detection time and sample rate, for a specific channel under certain requirements on efficiency and interference at a confidence level. These issues are critical in the design of a CR spectrum sensing system, which address fundamental impacts on the resultant system performance.

The rest of this paper is organized as follows. The model setting and the CFAR and CDR principles for energy detection are introduced in Section 2, where the thresholds are derived by assuming that all signals are Gaussian. The principle of compromise for CFAR and CDR is introduced in Section 3, and the two presentations of optimization are introduced and theoretically analyzed. Some numerical experiments are conducted to check the relevant evolutions of the solutions of the proposed optimization problems in Section 4. In Section 5, the fundamental limitations in energy detection are demonstrated, and some asymptotical orders of the critical values are discovered via theoretical analysis. Finally, the conclusive remarks of this study are given in Section 6.

## 2. Model Setting and Thresholds by CFAR and CDR

A block diagram of typical energy detection for spectrum sensing is shown in Figure 1. The input Band Pass Filter (BPF) which has bandwidth  $W$  centered at  $f_c$  aims to remove the out-of-band signals. Note that  $W$  and  $f_c$  must be known to the secondary user so that it can perform spectrum sensing for the corresponding channels. After the signal is digitized by an analog-to-digital converter (ADC), a simple square and average device is used to estimate the received signal energy. Assume the input signal to the energy detection is real. The estimated energy,  $u$ , is then compared with a threshold,  $\lambda$ , to decide if a signal is present ( $H_1$ ) or not ( $H_0$ ).

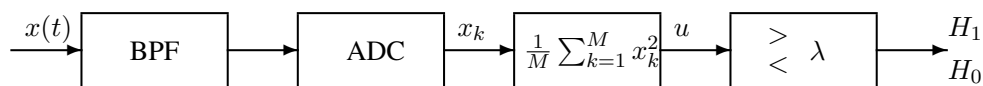


Figure 1. Block diagram of an energy detector.

Spectrum sensing is to determine whether a licensed band is currently used by its primary user. This can be formulated into a binary hypothesis testing problem [19,20]:

$$x(k) = \begin{cases} n(k), & H_0(\text{vacant}), \\ s(k) + n(k), & H_1(\text{occupied}), \end{cases} \quad (2)$$

where  $s(k)$ ,  $n(k)$ , and  $x(k)$  represent the primary user's signal, the noise, and the received signal, respectively. The noise is assumed to be an iid Gaussian random process of zero mean and variance  $\sigma_n^2$ , whereas the signal is also assumed to be an iid Gaussian random process of zero mean and variance of  $\sigma_s^2$ . The signal to noise ratio is defined as the ratio of signal variance to the noise variance

$$\text{SNR} = \sigma_s^2 / \sigma_n^2. \quad (3)$$

The test statistics generated from the energy detector, as shown in Figure 1, is

$$u = \frac{1}{M} \sum_{k=1}^M x_k^2. \quad (4)$$

The threshold is determined typically based on two standard principles: constant false alarm rate (CFAR) and constant detection rate (CDR) [14,15]. CDR guarantees a designated low level of interference to primary users, which nonetheless results in low spectrum usability of secondary users given a fixed sensing time. On the other hand, CFAR protects secondary users' spectrum utilization at a designated high level, which may lead to a high level of interference to primary users. Therefore, each of the CFAR and CDR principles can general ensure either one of the error probabilities under low level within a limited sensing time, i.e., false alarm probability  $P_{fa}$  for CFAR and missed detection probability  $(1 - P_d)$  for CDR, respectively.

Under Hypotheses  $H_0$  and  $H_1$ , the test statistics  $u$  is a random variable whose probability density function (PDF) is chi-square distributed. Let us denote a chi-square distributed random variable  $X$  with  $M$  degrees of freedom as  $X \sim \chi_M^2$ , and recall its PDF as

$$f_X(x, M) = \begin{cases} \frac{1}{2^{M/2}\Gamma(M/2)} x^{M/2-1} e^{-x/2}, & \text{for } x > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

where  $\Gamma(\cdot)$  denotes Gamma function, given in Equation (A2) in the Appendix.

Clearly, under Hypothesis  $H_0$ ,  $Mu/\sigma_n^2 \sim \chi_M^2$ ; and  $Mu/\sigma_t^2 \sim \chi_M^2$  under  $H_1$  with  $\sigma_t^2 = (1 + \text{SNR})\sigma_n^2$ . Thus, the PDF of test statistics  $u$ , given by test, is

$$f_u(x) \sim \begin{cases} \frac{\sigma_n^2}{M} f_X\left(\frac{x\sigma_n^2}{M}, M\right), & \text{under } H_0; \\ \frac{\sigma_t^2}{M} f_X\left(\frac{x\sigma_t^2}{M}, M\right), & \text{under } H_1. \end{cases} \quad (6)$$

Observe that  $Eu = \sigma_n^2$  and  $\text{var}(u) = \frac{2}{M}\sigma_n^4$  under  $H_0$ , by the central limit theorem [21], the test statistics  $u$  asymptotically obeys the Gaussian distribution with mean  $\sigma_n^2$  and variance  $\frac{2}{M}\sigma_n^4$ . Similar distribution can be derived under  $H_1$ . Therefore, when  $M$  is sufficiently large, we can approximate the PDF of  $u$  using a Gaussian distribution:

$$\tilde{f}_u(x) \sim \begin{cases} \mathcal{N}(\sigma_n^2, 2\sigma_n^4/M), & \text{under } H_0; \\ \mathcal{N}(\sigma_t^2, 2\sigma_t^4/M), & \text{under } H_1. \end{cases} \quad (7)$$

For a given threshold  $\lambda$ , the probability of false alarm is given by

$$P_{fa}(\lambda) = \text{prob}[u > \lambda | H_0] = \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_n^2}\right), \quad (8)$$

where  $\Gamma(a, x)$  is the upper incomplete gamma function in Equation (A3) in the Appendix, and the approximated form of  $P_{fa}$  corresponding to distribution in Equation (7) for large  $M$  is

$$\tilde{P}_{fa}(\lambda) = Q\left(\frac{\lambda - \sigma_n^2}{\sigma_n^2 / \sqrt{M/2}}\right), \quad (9)$$

where  $Q(\cdot)$  is defined in Equation (A1) in the Appendix.

In practice, if it is required to guarantee a reuse probability of the unused spectrum, the probability of false alarm is fixed to a small value (e.g.,  $P_{fa} = 0.05$ ) and meanwhile the detection probability is expected to be maximized as much as possible. This is referred to as constant false alarm rate (CFAR) principle [19,22]. Under the CFAR principle, the probability of false alarm rate ( $P_{fa}$ ) is predetermined, and the threshold ( $\lambda_{fa}$ ) can be set accordingly by

$$\lambda_{fa} = \frac{2\sigma_n^2}{M} \Gamma^{-1}\left(M/2, P_{fa}\right), \quad (10)$$

where  $\Gamma^{-1}(a, x)$  is the inverse function of  $\Gamma(a, x)$ . For the approximation case corresponding to distribution Equation (7) for large  $M$ , the threshold is

$$\tilde{\lambda}_{fa} = \sigma_n^2 \left(1 + \frac{Q^{-1}(P_{fa})}{\sqrt{M/2}}\right), \quad (11)$$

where  $Q^{-1}(x)$  is the inverse function of  $Q(x)$ .

Similarly, under Hypothesis  $H_1$ , for a given threshold  $\lambda$ , the probability of detection is given by

$$P_d(\lambda) = \text{prob}[u > \lambda | H_1] = \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_t^2}\right), \quad (12)$$

where  $\Gamma(a, x)$  is the upper incomplete gamma function. Its approximating form of  $P_d$  corresponding to distribution in Equation (7) for large  $M$  is

$$\tilde{P}_d(\lambda) = Q\left(\frac{\lambda - \sigma_t^2}{\sigma_t^2 / \sqrt{M/2}}\right). \quad (13)$$

Practically, if it is required to guarantee interference-free to the primary users, the probability of detection should be set high (e.g.,  $P_d = 0.95$ ) and the probability of false alarm should be minimized as much as possible. This is called the constant detection rate (CDR) principle [19,22]. With this, the threshold under the CDR principle to achieve a target probability of detection is given by

$$\lambda_d = \frac{2\sigma_n^2(1 + \text{SNR})}{M} \Gamma^{-1}\left(M/2, P_d\right). \quad (14)$$

The corresponding approximation case is

$$\tilde{\lambda}_d = \sigma_n^2(1 + \text{SNR}) \left(1 + \frac{Q^{-1}(P_d)}{\sqrt{M/2}}\right). \quad (15)$$

Due to the similarity of Equations (10) and (14), we can expect that the derivation of the threshold values for CFAR and CDR are similar. Thus, it is not surprising to see that some analytic results derived

by assuming CFAR based detection can be applied to CDR based detection with minor modifications and vice versa (see, e.g., [19,22]).

### 3. Thresholds by New Principle

It is clear that using CFAR and CDR principles can guarantee a low  $P_{fa}$  and  $(1 - P_d)$ , respectively. However, in practice, we may hope both of them to be low. This motivates us to come up with a new principle such that a threshold is determined to keep the sum of the two error probabilities at a designated low level. The problem of interest in the study is to find a threshold, for a given small  $\alpha \in (0, 1)$ , say  $\alpha = 0.05$ , such that

$$P_{fa}(\lambda) + (1 - P_d(\lambda)) \leq \alpha, \tag{16}$$

where  $P_{fa}(\lambda)$  and  $P_d(\lambda)$  are given by Equations (8) and (12), respectively. This is nonetheless not a well-posed presentation, since there may be too many solutions or no solution to the inequality in Equation (16) with respect to  $\lambda$  for a given noise variance  $\sigma_n^2$ , SNR (or  $\sigma_t^2$ ), and data size  $M$ . In this section, a suite of well-posed presentations for realizing this idea are formulated and analyzed, and relevant properties are developed for reference. Specifically, the presentations considered in the study include the following two scenarios:

- (i) By assuming given  $\sigma_n^2$  and SNR (or  $\sigma_t^2$ ), our target is to find minimum data size  $M$  and the corresponding threshold  $\lambda$  satisfying the inequality in Equation (16). This results in a nonlinear optimization problem as following:

$$(NP_1) \begin{cases} \text{Min} & M \\ \text{s.t.} & \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_n^2}\right) + 1 - \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_t^2}\right) \leq \alpha. \end{cases}$$

- (ii) By taking  $\sigma_n^2$  and  $M$  fixed, we target to find minimum SNR and corresponding threshold  $\lambda$  satisfying the inequality in Equation (16). This also results in a nonlinear optimization problem as following:

$$(NP_2) \begin{cases} \text{Min} & \text{SNR} \\ \text{s.t.} & \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_n^2}\right) + 1 - \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_t^2}\right) \leq \alpha. \end{cases}$$

We find in Proposition 1 that the threshold can be unambiguously determined if  $\sigma_n^2$  and SNR are given. Based on this theoretical discovery, the numerical algorithm for solving (NP<sub>1</sub>) and (NP<sub>2</sub>) can be significantly simplified.

**Proposition 1.** *Both in nonlinear optimization problems (NP<sub>1</sub>) and (NP<sub>2</sub>), if solvable, the solution for  $\lambda$  should be*

$$\lambda_0 = \frac{(1 + s)\sigma_n^2 \ln(1 + s)}{s}, \tag{17}$$

where  $s$  represents SNR for brief.

**Proof.** Let us consider the case for (NP<sub>1</sub>) only, since (NP<sub>2</sub>) has a very similar shape. We consider the Lagrange function of (NP<sub>1</sub>) with respect to a multiplier  $\mu$  as

$$L(M, \lambda, s_1, \mu) = M + \mu \left[ \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_n^2}\right) + 1 - \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_t^2}\right) + s_1 - \alpha \right], \tag{18}$$

where  $s_1 \geq 0$  is a slack variable. Clearly, differentiating  $L(M, \lambda, s, \mu)$  by  $\lambda$ , we have

$$\frac{\partial L(M, \lambda, s_1, \mu)}{\partial \lambda} = \frac{\mu}{\Gamma(\frac{M}{2})} \left[ \left( \frac{M\lambda}{2\sigma_t^2} \right)^{\frac{M}{2}-1} e^{-\frac{M\lambda}{2\sigma_t^2}} \frac{M}{2\sigma_t^2} - \left( \frac{M\lambda}{2\sigma_n^2} \right)^{\frac{M}{2}-1} e^{-\frac{M\lambda}{2\sigma_n^2}} \frac{M}{2\sigma_n^2} \right].$$

Let  $\frac{\partial L(M, \lambda, s_1, \mu)}{\partial \lambda} = 0$ , we get a simplified equivalent equation as

$$\frac{1}{\sigma_t^2} e^{-\frac{\lambda}{\sigma_t^2}} = \frac{1}{\sigma_n^2} e^{-\frac{\lambda}{\sigma_n^2}},$$

which further means

$$e^{\left(\frac{1}{\sigma_n^2} - \frac{1}{\sigma_t^2}\right)\lambda} = \frac{\sigma_t^2}{\sigma_n^2} = 1 + s.$$

Solving this equation, we derive Equation (17). □

The following proposition shows that the two nonlinear optimization problems (NP<sub>1</sub>) and (NP<sub>2</sub>) are well-posed, i.e., the solutions for (NP<sub>1</sub>) and (NP<sub>2</sub>) uniquely exist.

**Proposition 2.** Both nonlinear optimization problems (NP<sub>1</sub>) and (NP<sub>2</sub>) are well-posed: (i). for any given  $\alpha \in (0, 1)$ ,  $\sigma_n^2$  and  $SNR > 0$ , (NP<sub>1</sub>) has one and only one solution pair  $(M, \lambda)$ ; and (ii) for any given  $\alpha \in (0, 1)$ ,  $\sigma_n^2$  and  $M > 0$ , (NP<sub>2</sub>) has one and only one solution pair  $(SNR, \lambda)$ .

**Proof.** Let SNR be denoted by  $s$ , and the LHS of the restriction inequality of (NP<sub>1</sub>) be expressed by a function

$$\Gamma(M, s, \lambda) = \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_n^2}\right) + 1 - \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_t^2}\right). \tag{19}$$

By the definition of SNR, it follows that  $\sigma_t^2 = (1 + s)\sigma_n^2$ . Thus, we have

$$\frac{M\lambda}{2\sigma_n^2} - \frac{M\lambda}{2\sigma_t^2} = \frac{M\lambda\sigma_n^2}{2\sigma_t^2\sigma_n^2} = \frac{Ms\lambda}{2\sigma_t^2}. \tag{20}$$

Note that the threshold  $\lambda$  should be located between the two energies  $\sigma_n^2$  and  $\sigma_t^2$ , i.e.,  $\lambda \in (\sigma_n^2, \sigma_t^2)$ . We know that, for small SNR ( $s$ ) and small  $M$ , the distance between  $\frac{M\lambda}{2\sigma_n^2}$  and  $\frac{M\lambda}{2\sigma_t^2}$  is very close. Hence, the value of  $\Gamma(M, s, \lambda)$  given by Equation (19) is close to 1, which means the restriction  $\Gamma(M, s, \lambda) \leq \alpha$  is probably violated.

Next, we demonstrate that the solution of (NP<sub>1</sub>) is unique if it exists. Clearly, it is sufficient to show that  $\frac{\partial \Gamma(M, s, \lambda)}{\partial M} < 0$ . For this, by noticing that

$$\begin{aligned} \frac{\partial \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_n^2}\right)}{\partial M} &= -\frac{1}{\Gamma\left(\frac{M}{2}\right)} \frac{d\Gamma\left(\frac{M}{2}\right)}{dM} \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_n^2}\right) - \frac{1}{\Gamma\left(\frac{M}{2}\right)} \left(\frac{M\lambda}{2\sigma_n^2}\right)^{\frac{M}{2}-1} e^{-\frac{M\lambda}{2\sigma_n^2}} \frac{\lambda}{2\sigma_n^2} \\ &\quad + \frac{1}{\Gamma\left(\frac{M}{2}\right)} \int_{\frac{M\lambda}{2\sigma_n^2}}^{\infty} e^{-t} t^{\frac{M}{2}-1} \frac{\ln t}{2} dt \end{aligned} \tag{21}$$

and  $\frac{\partial \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_t^2}\right)}{\partial M}$ , similar to Equation (21) by replacing  $\sigma_n$  by  $\sigma_t$  therein, we have



$$\frac{\partial \Gamma(M, s, \lambda)}{\partial M} \Big|_{\lambda=\lambda_0} = \frac{1}{\Gamma(\frac{M}{2})} \int_{\frac{M\lambda_0}{2\sigma_t^2}}^{\frac{M\lambda_0}{2\sigma_n^2}} e^{-t} t^{\frac{M}{2}-1} \left[ \frac{1}{\Gamma(\frac{M}{2})} \frac{d\Gamma(\frac{M}{2})}{dM} - \frac{\ln t}{2} \right] dt \triangleq \frac{D(M, s)}{\Gamma(\frac{M}{2})}, \tag{22}$$

where  $\lambda_0 = \frac{(1+s)\sigma_n^2 \ln(1+s)}{s}$  is given by Equation (17).

Denote

$$a \triangleq \frac{M\lambda_0}{2\sigma_t^2} = \frac{M \ln(1+s)}{2s}, \quad b \triangleq \frac{M\lambda_0}{2\sigma_n^2} = \frac{M(1+s) \ln(1+s)}{2s}.$$

Clearly,  $b = a(1+s)$ . By noticing that  $a \xrightarrow{s \rightarrow 0} \frac{M}{2}$  and  $b \xrightarrow{s \rightarrow 0} \frac{M}{2}$ , we have  $\lim_{s \rightarrow 0} D(M, s) = 0$ . Similarly, by  $a \xrightarrow{s \rightarrow \infty} 0$  and  $b \xrightarrow{s \rightarrow \infty} \infty$ , it follows that  $\lim_{s \rightarrow \infty} D(M, s) = 0$ . Thus, to find the sign of Equation (22), we analyze the derivative of  $D(M, s)$  with respect to  $s$  below:

$$\begin{aligned} \frac{\partial D(M, s)}{\partial s} &= e^{-a} a^{\frac{M}{2}-1} \left[ e^{-as} (1+s)^{\frac{M}{2}-1} (C_M - \frac{1}{2} \ln b) \frac{\partial b}{\partial s} - (C_M - \frac{1}{2} \ln a) \frac{\partial a}{\partial s} \right] \\ &\triangleq e^{-a} a^{\frac{M}{2}-1} D_1(M, s), \end{aligned} \tag{23}$$

where  $C_M = \frac{1}{\Gamma(\frac{M}{2})} \frac{d\Gamma(\frac{M}{2})}{dM}$ . Note that

$$\begin{aligned} e^{-as} &= e^{-\frac{M \ln(1+s)}{2}} = (1+s)^{-\frac{M}{2}}, \\ \frac{\partial b}{\partial s} &= (1+s) \frac{\partial a}{\partial s} + a, \end{aligned}$$

we proceed the essential terms of Equation (23) as

$$\begin{aligned} D_1(M, s) &= \frac{1}{1+s} (C_M - \frac{1}{2} \ln b) \frac{\partial b}{\partial s} - (C_M - \frac{1}{2} \ln a) \frac{\partial a}{\partial s} \\ &= \frac{\partial a}{\partial s} \frac{1}{2} \ln \frac{a}{b} + \frac{a}{1+s} (C_M - \frac{1}{2} \ln b) \\ &= \frac{M}{2s(1+s)} \left[ \frac{\ln^2(1+s)}{2s} + (C_M - \frac{1}{2} \ln \frac{M}{2} - \frac{1}{2}) \ln(1+s) + \delta(s) \right], \end{aligned} \tag{24}$$

where  $\delta(s) = \frac{1}{2} \ln(1+s) \ln \frac{s}{\ln(1+s)}$ . Recalling an inequality

$$\ln x - \frac{1}{x} < \frac{\Gamma'(x)}{\Gamma(x)} < \ln x - \frac{1}{2x}$$

for  $x > 1$ , we find

$$-\frac{2}{M} < C_M - \frac{1}{2} \ln \frac{M}{2} < -\frac{1}{M}.$$

Thus, the sign of  $D_1(M, s)$  changes from negative to positive as  $s$  moves from 0 to  $\infty$ , and it also does for  $\frac{\partial D(M, s)}{\partial s}$ . Together with the two limitations of  $D(M, s)$ , we know that  $D(M, s)$  decreases from  $D(M, 0) = 0$  to negative minimum and then increase to  $D(M, \infty) = 0$ . Thus,  $\frac{\partial \Gamma(M, s, \lambda)}{\partial M} \Big|_{\lambda=\lambda_0} < 0$  for  $s \in (0, \infty)$ , which derives the uniqueness of the solution.

Let us recall some basic facts of Gamma distribution below before the deducing the existence of the solution of (NP<sub>1</sub>) when  $M$  is sufficiently large. For a Gamma distribution with density function as



$\frac{1}{\Gamma(k)} x^{k-1} e^{-x}$ , its expectation and deviation are  $k$  and  $\sqrt{k}$ , respectively. Let us first introduce a Gamma distributed random variable  $\xi$  with  $k = M/2$ . Denote

$$\begin{aligned} \frac{M\lambda}{2\sigma_n^2} - \frac{M}{2} &\triangleq \beta_1 \frac{M}{2}, \\ \frac{M}{2} - \frac{M\lambda}{2\sigma_t^2} &\triangleq \beta_2 \frac{M}{2}, \end{aligned}$$

where  $\beta_1 > 0$  and  $\beta_2 > 0$ , by the fact that  $\lambda \in (\sigma_n^2, \sigma_t^2)$ . By Equation (A3) (in the Appendix) and Chebyshev's inequality, we have

$$\begin{aligned} \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_n^2}\right) &= \text{P}\left[\xi - E\xi > \beta_1 \frac{M}{2}\right] \leq \text{P}\left[|\xi - E\xi| > \beta_1 \frac{M}{2}\right] \\ &\leq \frac{\text{Var}(\xi)}{\beta_1^2 (M/2)^2} = \frac{2}{\beta_1^2 M} \xrightarrow{M \rightarrow \infty} 0. \end{aligned} \tag{25}$$

Similarly,

$$\begin{aligned} 1 - \Gamma\left(\frac{M}{2}, \frac{M\lambda}{2\sigma_t^2}\right) &= \text{P}\left[E\xi - \xi > \beta_2 \frac{M}{2}\right] \leq \text{P}\left[|\xi - E\xi| > \beta_2 \frac{M}{2}\right] \\ &\leq \frac{\text{Var}(\xi)}{\beta_2^2 (M/2)^2} = \frac{2}{\beta_2^2 M} \xrightarrow{M \rightarrow \infty} 0. \end{aligned} \tag{26}$$

Hence, by Equation (19),  $\Gamma(M, s, \lambda) \rightarrow 0$  as  $M \rightarrow \infty$ . This means the restriction of (NP<sub>1</sub>) can be satisfied if  $M$  is sufficiently large, which derives Assertion (i).

For Assertion (ii), by Equation (20), we also know that the distance between  $\frac{M\lambda}{2\sigma_n^2}$  and  $\frac{M\lambda}{2\sigma_t^2}$  is very close for small SNR ( $s$ ). Thus, the restriction  $\Gamma(M, s, \lambda) \leq \alpha$  is probably violated.

Differentiating  $\Gamma(M, s, \lambda)$  by  $s$ , we have

$$\frac{\partial \Gamma(M, s, \lambda)}{\partial s} = \frac{-M\lambda}{2\Gamma(\frac{M}{2})} \left(\frac{M\lambda}{2\sigma_t^2}\right)^{\frac{M}{2}-1} e^{-\frac{M\lambda}{2\sigma_t^2} \frac{\sigma_n^2}{\sigma_t^2}} < 0.$$

Thus, we know that  $\Gamma(M, s, \lambda)$  decreases as  $s \rightarrow \infty$ . This means the solution is unique. By noticing that

$$\frac{M\lambda_0}{2\sigma_t^2} = \frac{M \ln(1+s)}{2s} \xrightarrow{s \rightarrow \infty} 0, \quad \frac{M\lambda_0}{2\sigma_n^2} = \frac{M(1+s) \ln(1+s)}{2s} \xrightarrow{s \rightarrow \infty} \infty,$$

which declares that  $\Gamma(M, s, \lambda) \rightarrow 0$  as  $s \rightarrow \infty$ . This further proves the existence of the solution, and, thus, Assertion (ii) follows.  $\square$

Similarly, by replacing  $P_{fa}$  and  $P_d$  in Equation (16) with the approximated distributions  $\tilde{P}_{fa}$  and  $\tilde{P}_d$  given by Equations (9) and (13), respectively, the corresponding case based on approximation distribution in Equation (7) for the new principle can be obtained. Precisely, for a given small  $\alpha > 0$ , a threshold  $\lambda$  can be identified such that

$$\tilde{P}_{fa}(\lambda) + (1 - \tilde{P}_d(\lambda)) \leq \alpha, \tag{27}$$

where  $\tilde{P}_{fa}(\lambda)$  and  $\tilde{P}_d(\lambda)$  are given by Equations (9) and (13), respectively. To achieve this, we have to formulate the two nonlinear optimization problems. Firstly, with given  $\sigma_n^2$  and SNR (or  $\sigma_t^2$ ), we try

to derive the minimum data size  $M$  and the corresponding threshold  $\lambda$  satisfying the inequality in Equation (27), the nonlinear optimization problem is formulated as:

$$(\widetilde{\text{NP}}_1) \begin{cases} \text{Min} & M \\ \text{s.t.} & Q\left(\frac{\lambda - \sigma_n^2}{\sigma_n^2/\sqrt{M/2}}\right) + 1 - Q\left(\frac{\lambda - \sigma_t^2}{\sigma_t^2/\sqrt{M/2}}\right) \leq \alpha. \end{cases}$$

Secondly, let  $\sigma_n^2$  and  $M$  be fixed, such that the minimum SNR and corresponding threshold  $\lambda$  satisfying the inequality in Equation (27) will be identified. This is also a nonlinear optimization problem as described as follows:

$$(\widetilde{\text{NP}}_2) \begin{cases} \text{Min} & \text{SNR} \\ \text{s.t.} & Q\left(\frac{\lambda - \sigma_n^2}{\sigma_n^2/\sqrt{M/2}}\right) + 1 - Q\left(\frac{\lambda - \sigma_t^2}{\sigma_t^2/\sqrt{M/2}}\right) \leq \alpha. \end{cases}$$

It is also found below that the potential threshold can be deterministically selected with given  $\sigma_n^2$  and SNR and data size  $M$ . By this theoretical discovery, the numerical algorithm to solve  $(\widetilde{\text{NP}}_1)$  and  $(\widetilde{\text{NP}}_2)$  can be largely simplified.

**Proposition 3.** In both nonlinear optimization problems  $(\widetilde{\text{NP}}_1)$  and  $(\widetilde{\text{NP}}_2)$ , if solvable, the solution for  $\lambda$  should be

$$\tilde{\lambda}_0 = \frac{\sigma_t^2 \sigma_n^2 (\sigma_t^2 - \sigma_n^2) + \sqrt{\Delta}}{\sigma_t^4 - \sigma_n^4}, \tag{28}$$

where

$$\Delta = \sigma_t^4 \sigma_n^4 (\sigma_t^2 - \sigma_n^2) [\sigma_t^2 - \sigma_n^2 + \delta (\sigma_t^2 + \sigma_n^2)] \tag{29}$$

with  $\delta = \frac{4}{M} \ln(1 + \text{SNR})$ . A simplified form is

$$\tilde{\lambda}_0 = \frac{\sigma_n^2 (1 + s) \left(1 + \sqrt{1 + \delta + \frac{2\delta}{s}}\right)}{2 + s}, \tag{30}$$

where  $s$  represents SNR. To assure  $\tilde{\lambda}_0 \in (\sigma_n^2, \sigma_t^2)$ , it requires that

$$M > \frac{4 \ln(1 + s)}{s^2}. \tag{31}$$

**Proof.** In the following, only the case of  $(\widetilde{\text{NP}}_1)$  is proven since the proof for  $(\widetilde{\text{NP}}_2)$  is similar. We first construct the Lagrange function for  $(\widetilde{\text{NP}}_1)$  with respect to a multiplier  $\mu$  as

$$\tilde{L}(M, \lambda, s_2, \mu) = M + \mu \left[ Q\left(\frac{\lambda - \sigma_n^2}{\sigma_n^2/m}\right) + 1 - Q\left(\frac{\lambda - \sigma_t^2}{\sigma_t^2/m}\right) + s_2 - \alpha \right], \tag{32}$$

where  $m = \sqrt{M/2}$  and  $s_2 \geq 0$  is a slack variable. Clearly, differentiating  $\tilde{L}(M, \lambda, s, \mu)$  in terms of  $\lambda$ , we have

$$\frac{\partial \tilde{L}(M, \lambda, s_2, \mu)}{\partial \lambda} = \frac{\mu}{\sqrt{2\pi}} \left[ \frac{m}{\sigma_t^2} e^{-\frac{m^2(\lambda - \sigma_t^2)^2}{2\sigma_t^4}} - \frac{m}{\sigma_n^2} e^{-\frac{m^2(\lambda - \sigma_n^2)^2}{2\sigma_n^4}} \right].$$

Let  $\frac{\partial \tilde{L}(M, \lambda, s_2, \mu)}{\partial \lambda} = 0$ , we get a simplified equivalent equation as

$$\frac{(\lambda - \sigma_n^2)^2}{\sigma_n^4} - \frac{(\lambda - \sigma_t^2)^2}{\sigma_t^4} = \frac{2}{m^2} \ln \left( \frac{\sigma_t^2}{\sigma_n^2} \right) = \frac{2}{m^2} \ln(1 + \text{SNR}) \triangleq \delta,$$

which further means

$$(\sigma_t^4 - \sigma_n^4)\lambda^2 - 2\sigma_t^2\sigma_n^2(\sigma_t^2 - \sigma_n^2)\lambda - \delta\sigma_t^4\sigma_n^4 = 0.$$

Thus,

$$\lambda = \frac{\sigma_t^2\sigma_n^2(\sigma_t^2 - \sigma_n^2) \pm \sqrt{\Delta}}{\sigma_t^4 - \sigma_n^4},$$

where  $\Delta$  is given by Equation (29). Note that  $\lambda \in (\sigma_n^2, \sigma_t^2)$ , i.e., the threshold should be located between the two expectations of the distributions under Hypotheses  $H_0$  and  $H_1$ , and

$$\sqrt{\Delta} \geq \sigma_t^2\sigma_n^2(\sigma_t^2 - \sigma_n^2),$$

thus we derive Equation (28). Equation (30) follows by substituting  $\sigma_t^2 = (1 + s)\sigma_n^2$  in Equation (28).

Clearly,  $\tilde{\lambda}_0 > \sigma_n^2$ . For  $\tilde{\lambda}_0 > \sigma_t^2$ , it is sufficient to require that

$$1 + \sqrt{1 + \delta + \frac{2\delta}{s}} < 2 + s,$$

which is equivalent to Equation (31).  $\square$

From the above, we have demonstrated that the two nonlinear optimization problems  $(\tilde{\text{NP}}_1)$  and  $(\tilde{\text{NP}}_2)$  are well-posed.

**Proposition 4.** Both nonlinear optimization problems  $(\tilde{\text{NP}}_1)$  and  $(\tilde{\text{NP}}_2)$  are well-posed: (i) for any given  $\alpha \in (0, 1)$ ,  $\sigma_n^2$  and  $\text{SNR} > 0$ ,  $(\tilde{\text{NP}}_1)$  has one and only one solution pair  $(M, \lambda)$ ; and (ii) for any given  $\alpha \in (0, 1)$ ,  $\sigma_n^2$  and  $M > 0$ ,  $(\tilde{\text{NP}}_2)$  has one and only one solution pair  $(\text{SNR}, \lambda)$ .

**Proof.** For convenience, let SNR be denoted by  $s$ , and the LHS of the restriction inequality of  $(\tilde{\text{NP}}_1)$  be expressed by a function

$$Q(M, s, \lambda) = Q\left(\frac{m(\lambda - \sigma_n^2)}{\sigma_n^2}\right) + 1 - Q\left(\frac{m(\lambda - \sigma_t^2)}{\sigma_t^2}\right), \tag{33}$$

where  $m = \sqrt{M/2}$ . For small SNR,  $\sigma_t^2$  and  $\sigma_n^2$  are very close. Thus, for small  $M$ , it is impossible to require that  $Q(M, s, \lambda) \leq \alpha$ , since  $\frac{m(\lambda - \sigma_n^2)}{\sigma_n^2}$  and  $\frac{m(\lambda - \sigma_t^2)}{\sigma_t^2}$  are too close.

(i) By differentiating  $Q(M, s, \lambda)$  in terms of  $m$ , we have

$$\frac{\partial Q(M, s, \lambda)}{\partial m} = \frac{1}{\sqrt{2\pi}} \left[ \frac{\lambda - \sigma_t^2}{\sigma_t^2} e^{-\frac{m^2(\lambda - \sigma_t^2)^2}{2\sigma_t^4}} - \frac{\lambda - \sigma_n^2}{\sigma_n^2} e^{-\frac{m^2(\lambda - \sigma_n^2)^2}{2\sigma_n^4}} \right] < 0,$$

by noticing that  $\lambda \in (\sigma_n^2, \sigma_t^2)$ . This concludes the uniqueness of the solution. Now, we derive the existence of the solution. Denote  $\tilde{\lambda}_0$  as the possible threshold of  $\lambda$  given by Equation (28). Clearly,

$$\tilde{\lambda}_0 \xrightarrow{m \rightarrow \infty} \frac{2\sigma_t^2\sigma_n^2}{\sigma_t^2 + \sigma_n^2} = \frac{2(1+s)\sigma_n^2}{2+s} \triangleq \lambda_*, \tag{34}$$

thus, for sufficiently large  $m$ , we have

$$\tilde{\lambda}_0 - \sigma_n^2 \geq \gamma_1|\lambda_* - \sigma_n^2|, \quad \sigma_t^2 - \tilde{\lambda}_0 \geq \gamma_2|\lambda_* - \sigma_t^2|,$$

where  $\gamma_1 > 0$  and  $\gamma_2 > 0$ . Thus, we further have

$$Q\left(\frac{m(\tilde{\lambda}_0 - \sigma_n^2)}{\sigma_n^2}\right) \leq Q\left(\frac{m\gamma_1|\lambda_* - \sigma_n^2|}{\sigma_n^2}\right) \xrightarrow{m \rightarrow \infty} 0,$$

$$1 - Q\left(\frac{m(\lambda - \sigma_t^2)}{\sigma_t^2}\right) \leq Q\left(\frac{m\gamma_2|\lambda_* - \sigma_t^2|}{\sigma_t^2}\right) \xrightarrow{m \rightarrow \infty} 0.$$

These mean  $Q(M, s, \lambda) \xrightarrow{m \rightarrow \infty} 0$ , which implies the existence of solution.

(ii) By differentiating  $Q(M, s, \lambda)$  in terms of  $s$ , we have

$$\frac{\partial Q(M, s, \lambda)}{\partial s} = \frac{-m\lambda\sigma_n^2}{\sigma_t^4\sqrt{2\pi}} e^{-\frac{m^2(\lambda - \sigma_t^2)^2}{2\sigma_t^4}} < 0,$$

which declares the uniqueness of solution for this case. Consider  $\tilde{\lambda}_0$  given by Equation (28) as  $s \rightarrow \infty$ ; clearly,  $\tilde{\lambda}_0 \xrightarrow{s \rightarrow \infty} \infty$  with the order  $\sqrt{\ln s}$ , and  $\sigma_t^2 \xrightarrow{s \rightarrow \infty} \infty$  with the order  $s$ . Thus, we have

$$\tilde{\lambda}_0 - \sigma_n^2 \xrightarrow{s \rightarrow \infty} \infty, \quad \sigma_t^2 - \tilde{\lambda}_0 \xrightarrow{s \rightarrow \infty} \infty,$$

which means  $Q(M, s, \lambda) \xrightarrow{s \rightarrow \infty} 0$ , guaranteeing the existence of solution for this case.  $\square$

Based on the above propositions, the proposed principles for threshold selection can be well defined provided data size  $M$  is sufficiently large and SNR is given, or SNR is sufficiently large and  $M$  is given. Hence, mathematically, some fundamental limitations are identified regarding data size  $M$  when SNR is given, as well as SNR when  $M$  is given.

#### 4. Numerical Experiments

Two kinds of simulations are conducted in this section. One is to observe the evolutions of the solutions of the above four nonlinear optimization problems, and the other is to investigate the accuracy of the two asymptotical equations: Equation (61) for both (NP<sub>1</sub>) and ( $\tilde{\text{NP}}_1$ ) as SNR tends to zero, and Equation (62) for both (NP<sub>2</sub>) and ( $\tilde{\text{NP}}_2$ ) as  $M$  tends to infinity. The two formulae are discovered by Theorems 1–4.

##### 4.1. Intuitive Sense of the Solutions for Relevant Optimization Problems

Specifically, we would like to check the evolutions of solutions  $M$  in (NP<sub>1</sub>) and ( $\tilde{\text{NP}}_1$ ) as SNR tends to 0, respectively; and the evolutions of solutions SNR in (NP<sub>2</sub>) and ( $\tilde{\text{NP}}_2$ ) as  $M$  tends to  $\infty$ , respectively.

The nonlinear optimization problems (NP<sub>1</sub>)(NP<sub>2</sub>) and ( $\tilde{\text{NP}}_1$ )( $\tilde{\text{NP}}_2$ ) can be solved directly by Matlab code /fmincon/, which is used to find minimum of an object function with nonlinear constraints. However, it is found to be quite unstable when using /fmincon/ to solve them directly, say (NP<sub>1</sub>), most likely due to the nonlinearity in the constraint of (NP<sub>1</sub>).

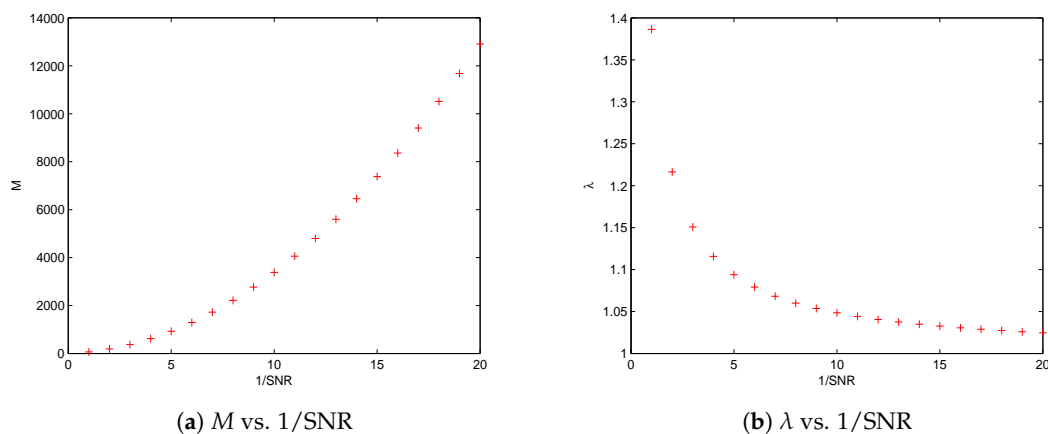
Proposition 1 significantly simplifies the solution process for (NP<sub>1</sub>) and meanwhile makes it stable by substituting the potential optimal threshold given by Equation (17) into the constraint of (NP<sub>1</sub>). The same for (NP<sub>2</sub>). As to ( $\tilde{\text{NP}}_1$ ) and ( $\tilde{\text{NP}}_2$ ), Proposition 3 plays the same role.

**Example 1.** Let  $\alpha = 0.05$  and  $\sigma_n = 1$ . We deal with  $(NP_1)$  by letting  $SNR = \frac{1}{k}$  with  $k = 1, 2, \dots, 20$ . The solutions for  $M$  and corresponding  $\lambda$  are shown in Figure 2a,b, respectively.

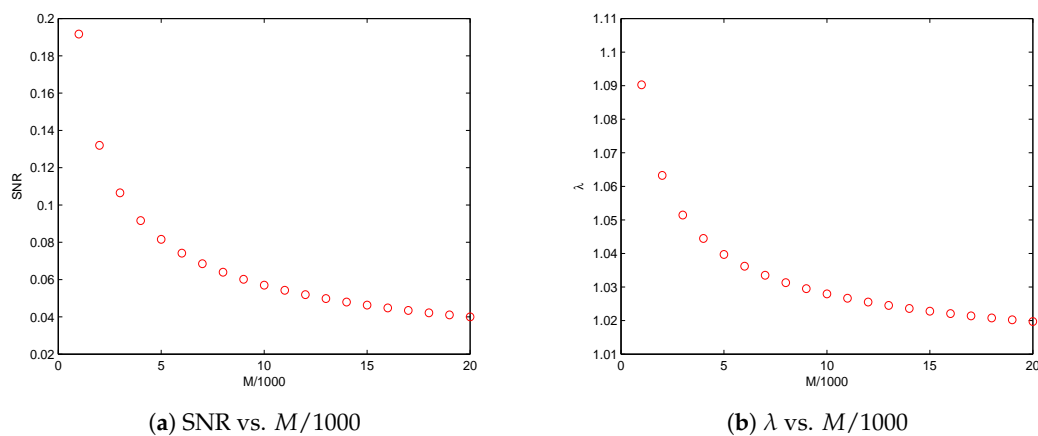
**Example 2.** Let  $\alpha = 0.05$  and  $\sigma_n = 1$ . We deal with  $(NP_2)$  by letting  $M = 1000k$  with  $k = 1, 2, \dots, 20$ . The solutions for  $SNR$  and  $\lambda$  are shown in Figure 3a,b, respectively.

**Example 3.** Let  $\alpha = 0.05$  and  $\sigma_n = 1$ . We deal with  $(\widetilde{NP}_1)$  by letting  $SNR = \frac{1}{k}$  with  $k = 1, 2, \dots, 20$ . Since the solutions for  $M$  and  $\lambda$  are very close to the counterparts of Example 1, we only plot the differences between the two examples (the quantities of Example 3 minus that of Example 1), shown practically negligible in Figure 4a,b, respectively.

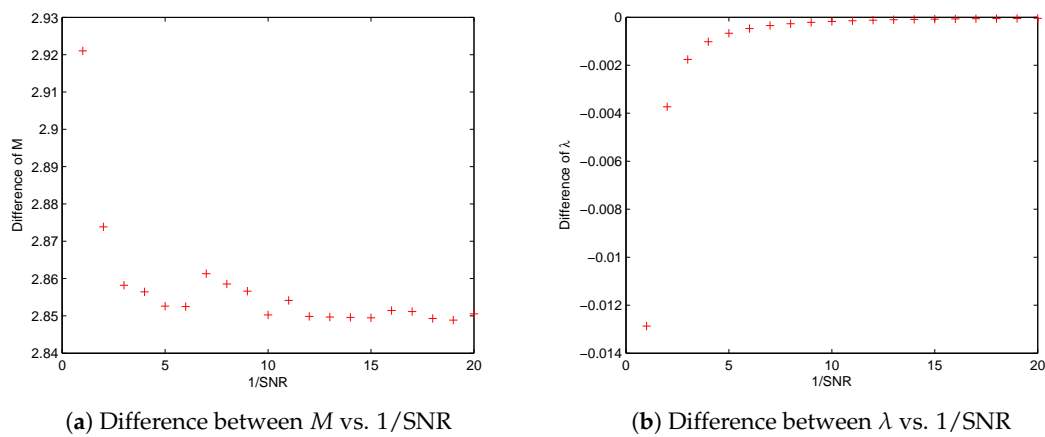
**Example 4.** Let  $\alpha = 0.05$  and  $\sigma_n = 1$ . We deal with  $(\widetilde{NP}_2)$  by letting  $M = 1000k$  with  $k = 1, 2, \dots, 20$ . Since the solutions for  $SNR$  and  $\lambda$  are very close to the counterparts of Example 2, we only plot the differences between the two examples (i.e., the quantities of Example 4 minus that of Example 2), as shown practically negligible in Figure 5a,b, respectively.



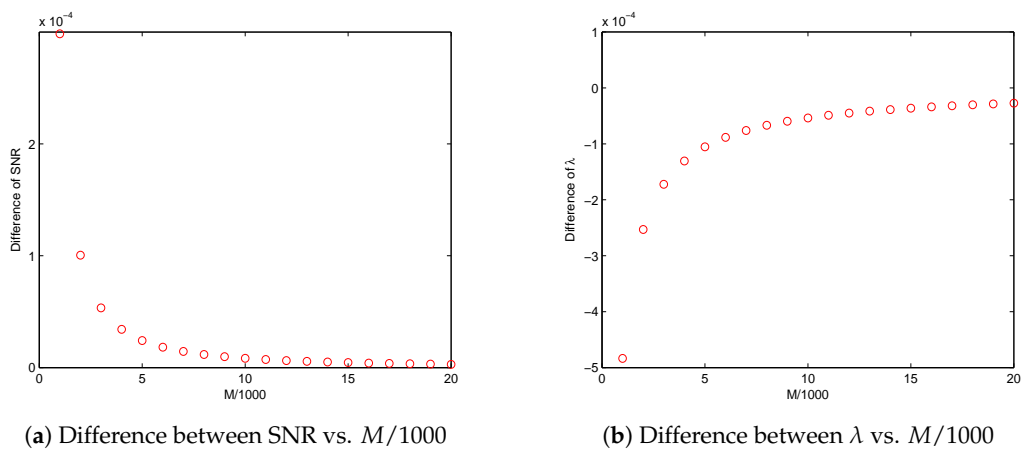
**Figure 2.** The plot of the solutions  $(M, \lambda)$  for  $(NP_1)$  with  $SNR = \frac{1}{k}$ ,  $k = 1, 2, \dots, 20$ , under setting  $\alpha = 0.05$  and  $\sigma_n = 1$ . (a) The plot of the solutions  $M$  of  $(NP_1)$  vs.  $1/SNR$ . (b) The plot of the corresponding  $\lambda$  vs.  $1/SNR$ .



**Figure 3.** The plot of solutions  $(SNR, \lambda)$  for  $(NP_2)$  with  $M = 1000k$  with  $k = 1, 2, \dots, 20$ ,  $k = 1, 2, \dots, 20$ , under setting  $\alpha = 0.05$  and  $\sigma_n = 1$ . (a) The plot of  $SNR$  vs.  $M/1000$ . (b) The plot of corresponding  $\lambda$  vs.  $M/1000$ .



**Figure 4.** The plot of differences between the solutions  $(M, \lambda)$  for  $(\widetilde{\text{NP}}_1)$  and  $(\text{NP}_1)$  with  $\text{SNR} = \frac{1}{k}$ ,  $k = 1, 2, \dots, 20$ , under setting  $\alpha = 0.05$  and  $\sigma_n = 1$ . (a) The plot of the differences between solutions  $M$  of  $(\widetilde{\text{NP}}_1)$  and  $(\text{NP}_1)$  vs.  $1/\text{SNR}$ . (b) The plot of differences for the corresponding  $\lambda$  vs.  $1/\text{SNR}$ .



**Figure 5.** The plot of differences between the solutions  $(\text{SNR}, \lambda)$  for  $(\widetilde{\text{NP}}_2)$  and  $(\text{NP}_2)$  with  $M = 1000k$ ,  $k = 1, 2, \dots, 20$ , under setting  $\alpha = 0.05$  and  $\sigma_n = 1$ . (a) The plot of the differences between solutions SNR of  $(\widetilde{\text{NP}}_2)$  and  $(\text{NP}_2)$  vs.  $M/1000$ . (b) The plot of differences for the corresponding  $\lambda$  vs.  $M/1000$ .

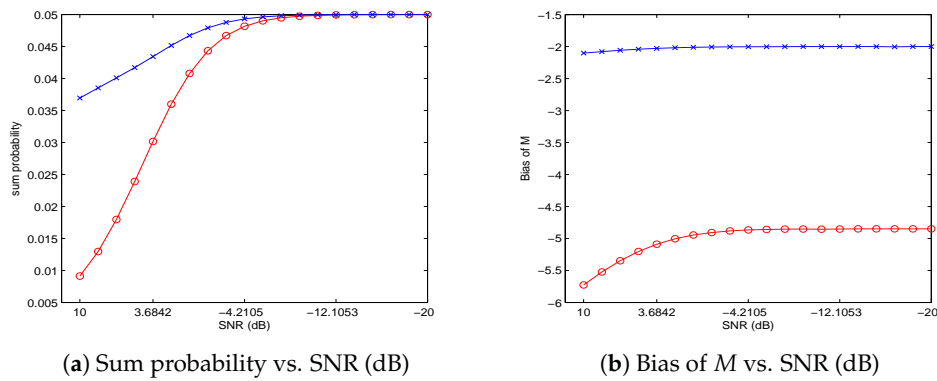
With Propositions 1 and 3, it is under our expectation that the thresholds  $\lambda$  tend to  $\sigma_n^2 = 1$  in Examples 1 and 3 as SNR tends to 0. The thresholds in Examples 2 and 4 tend to  $\sigma_n^2 = 1$  as  $M \rightarrow \infty$  due to the fact that the minimum SNR required tends to 0 as  $M \rightarrow \infty$ .

#### 4.2. Accuracy of the Two Asymptotical Equations (61) and (62)

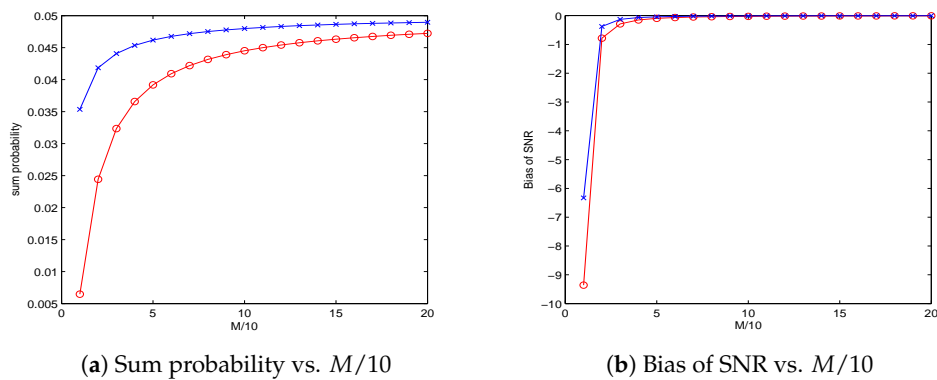
To investigate the performances of asymptotical Equation (61) for both  $(\text{NP}_1)$  and  $(\widetilde{\text{NP}}_1)$  as SNR tends to zero and Equation (62) for both  $(\text{NP}_2)$  and  $(\widetilde{\text{NP}}_2)$  as  $M$  tends to infinity, we design two more simulation examples below.

**Example 5.** Let  $\alpha = 0.05$  and  $\sigma_n = 1$ . For both  $(\text{NP}_1)$  and  $(\widetilde{\text{NP}}_1)$ , let  $\text{SNR} = 10^{1-3(k-1)/19}$ ,  $k = 1, 2, \dots, 20$ , corresponding to the range of 10 dB to  $-20$  dB. Let  $M$  be defined by Equation (61) and the corresponding  $\lambda$  is given by Equation (17) for  $(\text{NP}_1)$  and Equation (30) for  $(\widetilde{\text{NP}}_1)$ . The constraint summation probabilities of  $(\text{NP}_1)$  and  $(\widetilde{\text{NP}}_1)$ , marked by “-o” and “-x”, are plotted in Figure 6a. The bias between the solutions  $M$  from either  $(\text{NP}_1)$  or  $(\widetilde{\text{NP}}_1)$  and the asymptotical Equation (61) are plotted in Figure 6b. It is clear from both figures that the asymptotical Equation (61) provides quite accurate estimation of the solution for both  $(\text{NP}_1)$  and  $(\widetilde{\text{NP}}_1)$ , meanwhile, the required confidence level  $\alpha = 0.05$  is also satisfied.

**Example 6.** Let  $\alpha = 0.05$  and  $\sigma_n = 1$ . For both  $(NP_2)$  and  $(\widetilde{NP}_2)$ , let  $M = 10k$  with  $k = 1, 2, \dots, 20$ . Let SNR be defined by Equation (62) and the corresponding  $\lambda$  is given by Equation (17) for  $(NP_1)$  and Equation (30) for  $(\widetilde{NP}_1)$ . The constraint summation probabilities of  $(NP_2)$  and  $(\widetilde{NP}_2)$ , marked by “-o” and “-x”, are plotted in Figure 7a. The bias between the solutions SNR from either  $(NP_2)$  or  $(\widetilde{NP}_2)$  and the asymptotical Equation (62) are plotted in Figure 7b. It is clear from both figures that the asymptotical Equation (62) provides quite accurate estimation of the solution for both  $(NP_2)$  and  $(\widetilde{NP}_2)$ ; meanwhile, the required confidence level  $\alpha = 0.05$  is also satisfied.



**Figure 6.** The plot for Example 5 under the setting  $SNR = 10^{1-3(k-1)/19}$ ,  $k = 1, 2, \dots, 20$ , corresponding to the range of 10 dB to -20 dB,  $\alpha = 0.05$  and  $\sigma_n = 1$ . (a) The plot of the constraint summation probabilities of  $(NP_1)$  and  $(\widetilde{NP}_1)$ , marked by “-o” and “-x”, respectively. (b) The plot of the bias between the solution  $M$  from either  $(NP_1)$  or  $(\widetilde{NP}_1)$  and the asymptotical Equation (61), marked by “-o” and “-x”, respectively.



**Figure 7.** The plot for Example 6 under the setting  $M = 10k$ ,  $k = 1, 2, \dots, 20$ , under setting  $\alpha = 0.05$  and  $\sigma_n = 1$ . (a) The plot of the constraint summation probabilities of  $(NP_2)$  and  $(\widetilde{NP}_2)$ , marked by “-o” and “-x”, respectively. (b) The plot of the bias between the SNR solutions from either  $(NP_2)$  or  $(\widetilde{NP}_2)$  and the asymptotical Equation (62), marked by “-o” and “-x”, respectively.

From the simulation Examples 5 and 6, it is clear that the asymptotical Equations (61) and (62) provide quite accurate estimation of the solutions for the relevant four optimization problems with the constraint inequalities satisfied. The constraint inequality is tight for Equation (61) when SNR is sufficient small, and for Equation (62) when  $M$  is sufficient large. The relative difference between the two optimization frameworks is insignificant. This shows the potential application value of the two Equations (61) and (62).



### 5. Fundamental Limits of Detection

Let us denote  $M_{\min}$  as the solution of  $(NP_1)$  or  $(\widetilde{NP}_1)$ , and  $SNR_{\min}$  as the solution of  $(NP_2)$  or  $(\widetilde{NP}_2)$ , respectively. Obviously, for a fixed SNR, it is impossible to find a threshold  $\lambda$  satisfying inequality in Equation (16) or Equation (27) if the data size is smaller than  $M_{\min}$ . Equivalently, for a fixed data size  $M$ , it is not possible to find a threshold  $\lambda$  satisfying inequality in Equation (16) or Equation (27) if the SNR is smaller than  $SNR_{\min}$ . It can be observed that there exists some fundamental limitations in the effort of keeping the sum of the two error probabilities smaller than a designated confidence level. Theoretically, it is impossible to explicitly solve the four optimization problems introduced in Section 3. In this section, we investigate the asymptotical performances of the solutions to the four nonlinear optimization problems, i.e., to find the orders of  $M_{\min}$  as SNR tends to 0, and the orders of  $SNR_{\min}$  as  $M$  tends to  $\infty$ .

We first analyze  $(\widetilde{NP}_1)(\widetilde{NP}_2)$ , since Q function is much better than incomplete Gamma function, e.g., Q function possesses an interesting property as  $Q(x) = 1 - Q(-x)$ . Then, we establish a relation between incomplete Gamma function given by Equation (A3) (in the Appendix) and Q function in Lemma 1 to facilitate the investigations for  $(NP_1)(NP_2)$ . Throughout this section, SNR is usually simplified as  $s$  for brief. Two functions  $\psi(t)$  and  $\phi(t)$  are called to be of equivalent order, if

$$\lim_{t \rightarrow t_0} \frac{\psi(t)}{\phi(t)} = 1.$$

and are simply denoted below by  $\psi(t) \sim \phi(t)$  as  $t \rightarrow t_0$ .

The solution of  $(\widetilde{NP}_1)$  is discussed in the followings as a starting point.

**Theorem 1.** *The solution of nonlinear optimization problem  $(\widetilde{NP}_1)$  with  $\alpha \in (0, \frac{1}{2})$ , denoted by  $M_{\min}$ , has bounds given as*

$$\frac{2(2+s)^2}{s^2} x_1^2 < M_{\min} < 1 + \frac{2(2+s)^2}{s^2} x_2^2, \tag{35}$$

where  $x_1 = Q^{-1}\left(\frac{\alpha + \alpha_1}{2}\right)$  and  $x_2 = Q^{-1}\left(\frac{\alpha - \alpha_2}{2}\right)$  with  $\alpha_1 = \frac{s+2}{\sqrt{\pi M_{\min}}}$  and  $\alpha_2 = \frac{s+2}{\sqrt{\pi(M_{\min}-1)}}$ . Additionally,

$$M_{\min} \sim \frac{2(2+s)^2}{s^2} \left[ Q^{-1}\left(\frac{\alpha}{2}\right) \right]^2 \tag{36}$$

as  $s \rightarrow 0$ .

**Proof.** Recall the notations  $\tilde{\lambda}_0$  and  $\lambda_*$  given by Equations (30) and (34), respectively, we derive

$$\sigma_n^2 < \lambda_* < \tilde{\lambda}_0 < \sigma_t^2$$

for  $M > \frac{4 \ln(1+s)}{s^2}$ . It clearly indicates  $M_{\min} > \frac{4 \ln(1+s)}{s^2}$ . Otherwise, the minimum point  $\tilde{\lambda}_0$  of function  $Q(M, s, \lambda)$ , given by Equation (33), does not belong to  $(\sigma_n^2, \sigma_t^2)$ . Thus, the minimum value of  $Q(M, s, \lambda)$  over  $(\sigma_n^2, \sigma_t^2)$  attains at either  $\lambda = \sigma_n^2$  or  $\sigma_t^2$ , which means  $Q(M, s, \lambda) > \frac{1}{2} > \alpha$  over  $(\sigma_n^2, \sigma_t^2)$ , contradicting the definition of  $M_{\min}$ . Thus,  $M_{\min} \rightarrow \infty$  as  $s \rightarrow 0$ .

Clearly, we have

$$\tilde{\lambda}_0 - \lambda_* = \frac{(1+s)(\delta_1 - 1)}{2+s} \sigma_n^2, \tag{37}$$

where  $\delta_1 = \sqrt{1 + \delta + \frac{2\delta}{s}}$  and  $\delta = \frac{4}{M} \ln(1+s)$ .

Using standard Mean Value Theorem to  $Q(M, s, \lambda)$ , given by Equation (33), with respect to  $\lambda = \tilde{\lambda}_0, \lambda_*$ , we have

$$Q(M, s, \tilde{\lambda}_0) = Q(M, s, \lambda_*) + \left. \frac{\partial Q(M, s, \lambda)}{\partial \lambda} \right|_{\lambda=\xi} (\tilde{\lambda}_0 - \lambda_*), \tag{38}$$

where  $\xi \in (\lambda_*, \tilde{\lambda}_0)$ . Note that

$$\begin{aligned} \left| \left. \frac{\partial Q(M, s, \lambda)}{\partial \lambda} \right|_{\lambda=\xi} \right| &= \frac{m}{\sqrt{2\pi}} \left| \frac{1}{\sigma_t^2} e^{-\frac{m^2(\xi-\sigma_t^2)^2}{2\sigma_t^4}} - \frac{1}{\sigma_n^2} e^{-\frac{m^2(\xi-\sigma_n^2)^2}{2\sigma_n^4}} \right| \\ &< \frac{m}{\sqrt{2\pi}} \left( \frac{1}{\sigma_t^2} + \frac{1}{\sigma_n^2} \right) = \frac{m}{\sigma_n^2 \sqrt{2\pi}} \frac{2+s}{1+s}, \end{aligned}$$

and

$$\begin{aligned} \delta_1 - 1 &= \sqrt{1 + \delta + \frac{2\delta}{s}} - 1 < 1 + \frac{\delta}{2} + \frac{\delta}{s} - 1 = \frac{(s+2)\delta}{2s} \\ &= \frac{2(s+2) \ln(1+s)}{Ms} < \frac{2(s+2)}{M}, \end{aligned}$$

we have, by Equation (38),

$$|Q(M, s, \tilde{\lambda}_0) - Q(M, s, \lambda_*)| < \frac{m(\delta_1 - 1)}{\sqrt{2\pi}} < \frac{2m(s+2)}{M\sqrt{2\pi}} = \frac{s+2}{m\sqrt{2\pi}} \triangleq \alpha_1(m, s). \tag{39}$$

Clearly,  $\alpha_1(m, s) \rightarrow 0$  as  $m \rightarrow \infty$ .

Observing that  $Q(M, s, \lambda_*) = 2Q\left(\frac{ms}{2+s}\right)$ , and recalling  $M_{\min}$ , the solution of  $(\widetilde{NP}_1)$ , by Equation (39) we know that

$$2Q\left(\frac{m_0s}{2+s}\right) - \alpha_1(m_0, s) < Q(M_{\min}, s, \tilde{\lambda}_0) \leq \alpha, \tag{40}$$

$$2Q\left(\frac{m_1s}{2+s}\right) + \alpha_1(m_1, s) > Q(M_{\min} - 1, s, \tilde{\lambda}_0) > \alpha, \tag{41}$$

where  $m_0 = \sqrt{M_{\min}/2}$  and  $m_1 = \sqrt{(M_{\min} - 1)/2}$ .

By solving the inequality in Equation (40), we derive a lower bound of  $M_{\min}$  as

$$M_{\min} > \frac{2(2+s)^2}{s^2} x_1^2,$$

where  $x_1 = Q^{-1}\left(\frac{\alpha + \alpha_1(m_0, s)}{2}\right)$ . By solving the inequality in Equation (41), we derive an upper bound of  $M_{\min}$  as

$$M_{\min} < 1 + \frac{2(2+s)^2}{s^2} x_2^2,$$

where  $x_2 = Q^{-1}\left(\frac{\alpha - \alpha_1(m_1, s)}{2}\right)$ .

Recalling the fact that  $M_{\min} \rightarrow \infty$  as  $s \rightarrow 0$  guaranteed in the beginning of the proof, Equation (36) follows immediately.  $\square$

Now, based on the facts revealed in the above proof, the consideration of  $(\widetilde{NP}_2)$  is much simplified.

**Theorem 2.** The solution of nonlinear optimization problem  $(\widetilde{NP}_2)$  with  $\alpha \in (0, \frac{1}{2})$ , denoted by  $SNR_{min}$ , has bounds as following:

$$\frac{2Q^{-1}(\frac{\alpha+\alpha_1}{2})}{\sqrt{\frac{M}{2} - Q^{-1}(\frac{\alpha+\alpha_1}{2})}} < SNR_{min} < \epsilon + \frac{2Q^{-1}(\frac{\alpha-\alpha_2}{2})}{\sqrt{\frac{M}{2} - Q^{-1}(\frac{\alpha-\alpha_2}{2})}}, \tag{42}$$

where  $\alpha_1 = \frac{SNR_{min}+2}{\sqrt{\pi M}}$  and  $\alpha_2 = \frac{SNR_{min}-\epsilon+2}{\sqrt{\pi M}}$  with  $\forall \epsilon > 0$ . Additionally,

$$SNR_{min} \sim \frac{2Q^{-1}(\frac{\alpha}{2})}{\sqrt{\frac{M}{2} - Q^{-1}(\frac{\alpha}{2})}} \tag{43}$$

as  $M \rightarrow \infty$ .

**Proof.** Recalling the solution of  $(\widetilde{NP}_2)$   $SNR_{min}$ , denoted by  $s_m$  for brief, by Equation (39) we have

$$2Q\left(\frac{ms_m}{2 + s_m}\right) - \alpha_1(m, s_m) < Q(M, s_m, \tilde{\lambda}_0) \leq \alpha, \tag{44}$$

$$2Q\left(\frac{m(s_m - \epsilon)}{2 + (s_m - \epsilon)}\right) + \alpha_1(m, s_m - \epsilon) > Q(M, s_m - \epsilon, \tilde{\lambda}_0) > \alpha, \tag{45}$$

where  $\epsilon > 0$  and  $\alpha_1(m, s)$  is given by Equation (39). After solving the two inequalities and Equations (44) and (45), we can derive Equation (42).

Note that  $\alpha_1 \xrightarrow{M \rightarrow \infty} 0$  and  $\alpha_2 \xrightarrow{M \rightarrow \infty} 0$ , Equation (43) follows directly.  $\square$

As already mentioned, we establish a relation between Q function and incomplete Gamma function before the considerations of  $(NP_1)$  and  $(NP_2)$ .

**Lemma 1.** Extend the incomplete Gamma function given by Equation (A3) (in the Appendix) by letting  $\Gamma(k, x) = 1$  for  $x < 0$ . Then, for  $\forall x$ ,

$$\left| \Gamma\left(\frac{M}{2}, \frac{M}{2} + x\sqrt{\frac{M}{2}}\right) - Q(x) \right| \leq \frac{14C}{\sqrt{2M}}, \tag{46}$$

where  $C = 0.7056$ .

**Proof.** Let us first recall a simple version of the Berry–Esseen inequality (see, e.g., page 670 of [23]) here: Let  $\xi_1, \xi_2, \dots$ , be iid random variables with  $E\xi_1 = \mu$ ,  $E\xi_1^2 = \sigma^2$ , and  $E|\xi_1 - E\xi_1|^3 = \rho < \infty$ . In addition, let  $S_n = \sum_{k=1}^n \xi_k$ . Let the distribution function of

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n(\sigma^2 - \mu^2)}}$$

be denoted by  $F_n(x)$ , and the normal distribution function be denoted by  $\Phi(x)$ . There exists a positive constant C such that, for all x and n,

$$|F_n(x) - \Phi(x)| \leq \frac{C\rho}{\sqrt{n(\sigma^2 - \mu^2)^{3/2}}}.$$

The best current bound for C was discovered by Shevtsova as 0.7056 [24] in 2007.

For M iid standardly normal distributed random variables  $\eta_k, k = 1, 2, \dots, M$ , we know that  $E\eta_k^2 = 1, E\eta_k^4 = 3, E\eta_k^6 = 15$ . On the other hand, by the definition of chi-square distribution, it is clear that  $S_M = \sum_{k=1}^M \eta_k^2$  is chi-square distributed with degree M, i.e.,  $S_M \sim \chi_M^2$ . Apply the Berry–Esseen

inequality to  $\xi_k = \eta_k^2, k = 1, \dots, M$ , by noticing  $\mu = 1, \sigma^2 = 3$  and  $\rho \leq 28$  (by the fact  $|\eta_k^2 - 1| \leq \eta_k^2 + 1$ ) for this case, we have

$$\left| \Pr \left[ \frac{S_M - M}{\sqrt{2M}} > x \right] - Q(x) \right| \leq \frac{14C}{\sqrt{2M}}, \tag{47}$$

where  $C = 0.7056$ .

Recalling the PDF of chi-square distribution given by Equation (5), we have

$$\begin{aligned} \Pr \left[ \frac{S_M - M}{\sqrt{2M}} > x \right] &= \Pr \left[ S_M > M + x\sqrt{2M} \right] \\ &= \int_{M+x\sqrt{2M}}^{\infty} \frac{t^{\frac{M}{2}-1} e^{-\frac{t}{2}} dt}{2^{\frac{M}{2}} \Gamma(\frac{M}{2})} \\ &= \int_{\frac{M}{2}+x\sqrt{\frac{M}{2}}}^{\infty} \frac{u^{\frac{M}{2}-1} e^{-u} du}{\Gamma(\frac{M}{2})} \\ &= \Gamma \left( \frac{M}{2}, \frac{M}{2} + x\sqrt{\frac{M}{2}} \right). \end{aligned} \tag{48}$$

Combining Equations (47) and (48), we derive Equation (46).  $\square$

Now, we are in the position to analyze the situation of  $(NP_1)$ .

**Theorem 3.** *The solution of nonlinear optimization problem  $(NP_1)$  with  $\alpha \in (0, 1)$ , denoted by  $M_{min}$ , has bounds as following:*

$$\frac{2(2+s)^2}{s^2} x_1^2 < M_{min} < 1 + \frac{2(2+s)^2}{s^2} x_2^2, \tag{49}$$

where  $x_1$  and  $x_2$  are given in the following proof. Additionally,

$$M_{min} \sim \frac{2(2+s)^2}{s^2} \left[ Q^{-1} \left( \frac{\alpha}{2} \right) \right]^2 \tag{50}$$

as  $s \rightarrow 0$ .

**Proof.** Recalling the functions  $\Gamma(M, s, \lambda)$  and  $Q(M, s, \lambda)$  given by Equations (19) and (33), respectively, by Equation (46), we have

$$|\Gamma(M, s, \lambda) - Q(M, s, \lambda)| \leq \frac{14C}{\sqrt{2M}}, \tag{51}$$

where  $C = 0.7056$ .

Recalling also  $\lambda_0$  and  $\lambda_*$  given by Equations (17) and (34), respectively, we have

$$\sigma_n^2 < \lambda_* < \lambda_0 < \sigma_t^2$$

by the fact  $\frac{2s}{2+s} < \ln(1+s)$  for  $s > 0$ , and by the inequality  $\ln(1+s) < s - \frac{s^2}{2} + \frac{s^3}{3}$  for  $s > 0$ , we further have

$$0 < \lambda_0 - \lambda_* = \frac{(1+s)\sigma_n^2}{s} \left[ \ln(1+s) - \frac{2s}{2+s} \right] < \frac{s^2(1+s)(1+2s)\sigma_n^2}{6(2+s)}. \tag{52}$$

Similar to the derivation of Equation (39), by Equation (52), we have

$$|Q(M, s, \lambda_0) - Q(M, s, \lambda_*)| < \frac{ms^2(1 + 2s)}{6\sqrt{2\pi}}. \tag{53}$$

Observing that  $Q(M, s, \lambda_*) = 2Q\left(\frac{ms}{2+s}\right)$ , and recalling  $M_{\min}$ , the solution of (NP<sub>1</sub>), by Equations (51) and (53) we know that

$$2Q\left(\frac{m_0s}{2+s}\right) - \alpha_2(m_0, s) < \Gamma(M_{\min}, s, \tilde{\lambda}_0) \leq \alpha, \tag{54}$$

$$2Q\left(\frac{m_1s}{2+s}\right) + \alpha_2(m_1, s) > \Gamma(M_{\min} - 1, s, \tilde{\lambda}_0) > \alpha, \tag{55}$$

where  $m_0 = \sqrt{M_{\min}/2}$ ,  $m_1 = \sqrt{(M_{\min} - 1)/2}$ , and

$$\alpha_2(m, s) = \frac{14C}{\sqrt{2M}} + \frac{ms^2(1 + 2s)}{6\sqrt{2\pi}}. \tag{56}$$

By solving the inequality in Equation (54), we derive a lower bound of  $M_{\min}$  as

$$M_{\min} > \frac{2(2 + s)^2}{s^2} x_1^2,$$

where  $x_1 = Q^{-1}\left(\frac{\alpha + \alpha_2(m_0, s)}{2}\right)$ . While by solving the inequality in Equation (55), we derive an upper bound of  $M_{\min}$  as

$$M_{\min} < 1 + \frac{2(2 + s)^2}{s^2} x_2^2,$$

where  $x_2 = Q^{-1}\left(\frac{\alpha - \alpha_2(m_1, s)}{2}\right)$ .

It is sufficient to prove Equation (50) based on the facts that  $\alpha_2(m_0, s) \rightarrow 0$  and  $\alpha_2(m_1, s) \rightarrow 0$  as  $s \rightarrow 0$ . For this, we first note that

$$0 < \frac{M\lambda_0}{2\sigma_n^2} - \frac{M\lambda_0}{2\sigma_f^2} = \frac{M \ln(1 + s)}{2} < \frac{Ms}{2} = m^2s.$$

If  $m^2s \xrightarrow{s \rightarrow 0} 0$ , then  $\Gamma(M, s, \lambda) \xrightarrow{s \rightarrow 0} 1$ . This means that  $M \xrightarrow{s \rightarrow 0} \infty$ . On the other side, we have

$$\begin{aligned} \frac{M\lambda_0}{2\sigma_n^2} - \frac{M}{2} &= \frac{M}{2} \left[ \frac{(1 + s) \ln(1 + s)}{s} - 1 \right] > \frac{m^2s}{2 + s}, \\ \frac{M}{2} - \frac{M\lambda_0}{2\sigma_f^2} &= \frac{M}{2} \left[ 1 - \frac{\ln(1 + s)}{s} \right] > \frac{m^2s(3 - 2s)}{6} \end{aligned}$$

by inequalities  $\ln(1 + s) > \frac{2s}{2+s}$  and  $\ln(1 + s) < s - \frac{s^2}{2} + \frac{s^3}{3}$  for  $s > 0$ . Thus, we know that  $m_0s = O(1)$  as  $s \rightarrow 0$ . Otherwise, by Chebyshev’s inequality, similar to Equations(25) and (26), we have  $\Gamma(M_{\min}, s, \lambda) \xrightarrow{s \rightarrow \infty} 0$ , which contradicts the fact that  $\Gamma(M_{\min}, s, \lambda)$  should be around the quantity  $\alpha > 0$ . Hence, the conclusions that  $\alpha_2(m_0, s) \xrightarrow{s \rightarrow \infty} 0$  and  $\alpha_2(m_1, s) \xrightarrow{s \rightarrow \infty} 0$  follow directly.  $\square$

Based on the foundation in above proof, the analysis of (NP<sub>2</sub>) is simplified as below.

**Theorem 4.** The solution of nonlinear optimization problem (NP<sub>2</sub>) with  $\alpha \in (0, 1)$ , denoted by  $SNR_{min}$ , has bounds as follows:

$$\frac{2Q^{-1}\left(\frac{\alpha+\alpha_1}{2}\right)}{\sqrt{\frac{M}{2} - Q^{-1}\left(\frac{\alpha+\alpha_1}{2}\right)}} < SNR_{min} < \epsilon + \frac{2Q^{-1}\left(\frac{\alpha-\alpha_2}{2}\right)}{\sqrt{\frac{M}{2} - Q^{-1}\left(\frac{\alpha-\alpha_2}{2}\right)}}, \tag{57}$$

where  $\alpha_1$  and  $\alpha_2$  are given in the following proof, equipped with  $\forall \epsilon > 0$ . Additionally,

$$SNR_{min} \sim \frac{2Q^{-1}\left(\frac{\alpha}{2}\right)}{\sqrt{\frac{M}{2} - Q^{-1}\left(\frac{\alpha}{2}\right)}} \tag{58}$$

as  $M \rightarrow \infty$ .

**Proof.** Recalling the solution of (NP<sub>2</sub>)  $SNR_{min}$ , denoted by  $s_m$  for brief, and observing  $Q(M, s, \lambda_*) = 2Q\left(\frac{ms}{2+s}\right)$ , by Equations (51) and (53), we have

$$2Q\left(\frac{ms_m}{2+s_m}\right) - \alpha_2(m, s_m) < Q(M, s_m, \tilde{\lambda}_0) \leq \alpha, \tag{59}$$

$$2Q\left(\frac{m(s_m - \epsilon)}{2+(s_m - \epsilon)}\right) + \alpha_2(m, s_m - \epsilon) > Q(M, s_m - \epsilon, \tilde{\lambda}_0) > \alpha, \tag{60}$$

where  $\epsilon > 0$  and  $\alpha_2(m, s)$  is defined by Equation (56). Solving the two inequalities in Equations (59) and (60), we derive Equation (57) with  $\alpha_1 = \alpha_2(m, s_m)$  and  $\alpha_2 = \alpha_2(m, s_m - \epsilon)$ .

By the facts pointed out in last part of proof for Theorem 3, i.e., that  $m^2s$  is not an infinitesimal quantity (thus,  $s \rightarrow 0$  if  $m \rightarrow \infty$ ) and that  $ms$  is bounded, and recalling the definition in Equation (56) we have  $\alpha_1 \xrightarrow{M \rightarrow \infty} 0$  and  $\alpha_2 \xrightarrow{M \rightarrow \infty} 0$ . Thus, Equation (58) follows directly.  $\square$

Now, we find  $M_{min}$ , either the solution of (NP<sub>1</sub>) or ( $\widetilde{NP}_1$ ), has an asymptotical order as

$$M_{min} \sim \frac{2(2+s)^2}{s^2} \left[Q^{-1}\left(\frac{\alpha}{2}\right)\right]^2 \quad (s \rightarrow 0); \tag{61}$$

and  $SNR_{min}$ , either the solution of (NP<sub>2</sub>) or ( $\widetilde{NP}_2$ ), has an asymptotical order as

$$SNR_{min} \sim \frac{2Q^{-1}\left(\frac{\alpha}{2}\right)}{\sqrt{\frac{M}{2} - Q^{-1}\left(\frac{\alpha}{2}\right)}} \quad (M \rightarrow \infty). \tag{62}$$

If replacing the notations “ $\sim$ ” by “ $=$ ” in the above two formulas, we find that the derived two equations are equivalent by ignoring the differences between  $M$ , SNR and  $M_{min}$   $SNR_{min}$  respectively.

Based on the foundation in above proof, the analysis of (NP<sub>2</sub>) is simplified as below.

By noticing the facts that the potential threshold  $\tilde{\lambda}_0$  tends to  $\lambda_*$  as  $SNR \rightarrow 0$  (or  $M \rightarrow \infty$ ), and that  $Q(M, s, \lambda_*) = 2Q\left(\frac{ms}{2+s}\right)$ , we can propose another principle by replacing the restriction in Equation (27) in ( $\widetilde{NP}_1$ )( $\widetilde{NP}_2$ ) as

$$\tilde{P}_{fa}(\lambda_*) + (1 - \tilde{P}_d(\lambda_*)) \leq \alpha,$$

which is equivalent to

$$2Q\left(\frac{ms}{2+s}\right) \leq \alpha.$$

Under this principle, the corresponding  $(\widetilde{NP}_1)(\widetilde{NP}_2)$  issues can be explicitly solved. Such an idea of replacement also serves as the key for the proofs presented in this section.

## 6. Conclusions

Spectrum sensing is a key step of enabling the recently emerged CR technologies by detecting the presence/absence of signals to explore spatial and temporal availability of spectrum resources. Among the possible methods for spectrum sensing, energy detection is the most popular and widely adopted technique, most likely due to its low implementation complexity. Two detection principles, i.e., CFAR and CDR, have been reported to set a threshold for corresponding binary hypothesis. CDR protects primary users at a designated low level of interference, while CFAR ensures a high resource utilization available to the secondary users. In practice, it is desired to initiate a graceful tradeoff between these two principles.

Motivated by this, the paper explored a new principle where the sum of the false alarm probability  $P_{fa}$  from CFAR and the false detection probability  $(1 - P_d)$  from CDR is kept smaller than a predetermined confidence level. Mathematically, for a given small confidence level  $\alpha \in (0, 1)$ , say  $\alpha = 0.05$ , the proposed principle aims to identify a threshold  $\lambda$  such that

$$P_{fa}(\lambda) + (1 - P_d(\lambda)) \leq \alpha. \quad (63)$$

However, this equation regarding potential threshold  $\lambda$  may lead to too many solutions or no solutions for a given noise variance  $\sigma_n^2$ , SNR and data size  $M$ . To tackle this situation, the paper firstly introduced two well-posed presentations for the optimization problems by finding the minimum data size  $M$  (with  $\sigma_n^2$  and SNR given) and SNR (with  $\sigma_n^2$  and  $M$  given), respectively.

From our analysis, we found that for a fixed small SNR the data size  $M$  should be larger than a critical value, denoted by  $M_{\min}$ , to guarantee the existence of threshold  $\lambda$  suggested by the new principle under given confidence level  $\alpha$ . An asymptotical explicit form between  $M_{\min}$  and SNR, i.e., Equation (61), is further given in Section 5. On the other hand, it is also discovered that, for a given data size  $M$ , SNR should be greater than a minimum SNR to ensure the existence of threshold  $\lambda$  suggested by the new principle under a given confidence level  $\alpha$ . An asymptotical explicit form between the minimum SNR and  $M$ , i.e., Equations (62), is further proposed in Section 5. We found that, if data size  $M$  is fixed, SNR should be greater than a certain level to perform considerate detection. If SNR known to be small and fixed, the data size should be greater than a certain level to detect reliably. These discoveries are important for policymaking for the settings of a CR sensing system, such as detection time and design sample rate for a special channel to achieve efficiency and noninterference at a confidence level. The proposed optimization problems can be effectively solved fast by setting the initial solution given by Equation (61) or (62). Therefore, the proposed framework is applicable both in theoretical and operational aspects.

It is worth noting that the inequality in Equation (63) can be extended to a more general form as

$$w_1 P_{fa}(\lambda) + w_2 (1 - P_d(\lambda)) \leq \alpha, \quad (64)$$

where  $w_1 \geq 0$  and  $w_2 \geq 0$ . Clearly, if  $w_1 = 1$  and  $w_2 = 0$ , Equation (64) derives CFAR principle; if  $w_1 = 0$  and  $w_2 = 1$ , Equation (64) leads to CDR principle; and, finally, if  $w_1 = 1$  and  $w_2 = 1$ , Equation (64) turns to be the new principle introduced in this paper. It is of interest to consider relevant theories based on the inequality in Equation (64) in the general setting.

**Author Contributions:** Xiao-Li Hu conceived, analyzed, and wrote the original proposed framework under the supervision of Pin-Han Ho during a visit to Waterloo. Pin-Han Ho and Limei Peng gave advice on this work and helped revise the paper.

**Funding:** This research was funded by [National Research Foundation of Korea] grant number [2018R1D1A1B07051118].

**Conflicts of Interest:** The authors declare no conflict of interest.



## Appendix

The Appendix provides some preliminaries on a number of commonly used distributions in the paper for easy reading.

The complement of the standard normal distribution function is often denoted  $Q(x)$ , i.e.,

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt, \quad (\text{A1})$$

and is sometimes referred to simply as the Q-function, especially in engineering texts. This represents the tail probability of the standard Gaussian distribution. The Gamma function and regularized upper incomplete Gamma function are defined as

$$\Gamma(k) = \int_0^{\infty} t^{k-1} e^{-t} dt, \quad (\text{A2})$$

$$\Gamma(k, x) = \frac{1}{\Gamma(k)} \int_x^{\infty} t^{k-1} e^{-t} dt \quad (\text{A3})$$

for  $k > 0$ , respectively. We simply use  $Q^{-1}(x)$ ,  $\Gamma^{-1}(k, x)$  to present the inverse functions of  $Q(x)$ ,  $\Gamma(k, x)$  respectively.

## References

1. Liang, Y.-C.; Chen, K.-C.; Li, G.Y.; Mahonen, P. Cognitive radio networking and communications: An overview. *IEEE Trans. Veh. Technol.* **2011**, *60*, 3386–3407. [[CrossRef](#)]
2. Sahai, A.; Tandra, R.; Mishra, S.M.; Hoven, N. Fundamental design tradeoffs in cognitive radio systems. In Proceedings of the international Workshop on Technology and Policy for Accessing Spectrum, Boston, MA, USA, 5 August 2006.
3. Rahimzadeh, F.; Shahtalebi, K.; Parvaresh, F. Using NLMS algorithms in cyclostationary-based spectrum sensing for cognitive radio networks. *Wirel. Pers. Commun.* **2017**, *97*, 2781–2797. [[CrossRef](#)]
4. Deepa, B.; Iyer, A.P.; Murthy, C.R. Cyclostationary-based architecture for spectrum sensing in IEEE 802.22 WRAN. In Proceedings of the IEEE Global Telecommunications Conference GLOBECOM 2010, Miami, FL, USA, 6–10 December 2010.
5. Tsinos, C.G.; Berberidis, K. Adaptive eigenvalue-based spectrum sensing for multi-antenna cognitive radio systems. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 4454–4458.
6. Tsinos, C.G.; Berberidis, K. Decentralized adaptive eigenvalue-based spectrum sensing for multiantenna cognitive radio systems. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 1703–1715. [[CrossRef](#)]
7. Herath, S.P.; Rajatheva, N.; Tellambura, C. Energy detection of unknown signals in fading and diversity reception. *IEEE Trans. Commun.* **2011**, *59*, 2443–2453. [[CrossRef](#)]
8. Sofotasios, P.C.; Mohjazi, L.; Muhaidat, S.; Al-Qutayri, M.; Karagiannidis, G.K. Energy detection of unknown signals over cascaded fading channels. *IEEE Antennas Wirel. Propag. Lett.* **2016**, *15*, 135–138. [[CrossRef](#)]
9. Bagheri, A.; Sofotasios, P.C.; Tsiftsis, T.A.; Ho-Van, K.; Loupis, M.I.; Freear, S.; Valkama, M. Energy detection based spectrum sensing over enriched multipath fading channels. In Proceedings of the 2016 IEEE Wireless Communications and Networking Conference, Doha, Qatar, 3–6 April 2016.
10. Politis, C.; Maleki, S.; Tsinos, C.G.; Liolis, K.P.; Chatzinotas, S.; Ottersten, B. Simultaneous sensing and transmission for cognitive radios with imperfect signal cancellation. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 5599–5615. [[CrossRef](#)]
11. Politis, C.; Maleki, S.; Tsinos, C.; Chatzinotas, S.; Ottersten, B. On-board the satellite interference detection with imperfect signal cancellation. In Proceedings of the 2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Edinburgh, UK, 3–6 July 2016; pp. 1–5.
12. Ye, F.; Zhang, X.; Li, Y.; Tang, C. Faithworthy collaborative spectrum sensing based on credibility and evidence theory for cognitive radio networks. *Symmetry* **2017**, *9*, 36. [[CrossRef](#)]

13. Liu, P.; Qi, W.; Yuan, E.; Wei, L.; Zhao, Y. Full-duplex cooperative sensing for spectrum-heterogeneous cognitive radio networks. *Sensors* **2017**, *17*, 1773. [[CrossRef](#)] [[PubMed](#)]
14. Abraham, D.A. Performance analysis of constant-false-alarm-rate detectors using characteristic functions. *IEEE J. Ocean. Eng.* **2017**. [[CrossRef](#)]
15. Weinberg, G.V. Constant false alarm rate detection in Pareto Type II Clutter. *Digit. Signal Process.* **2017**, *68*, 192–198. [[CrossRef](#)]
16. Tandra, R.; Sahai, A. SNR walls for signal detection. *IEEE J. Sel. Top. Signal Process.* **2008**, *2*, 4–17. [[CrossRef](#)]
17. Carlos, K.C.; Birru, D. IEEE 802.22: An introduction to the first wireless standard based on cognitive radios. *J. Commun.* **2006**, *1*, 38–47.
18. Penna, F.; Pastrone, C.; Spirito, M.; Garello, R. An experimental study on spectrum sensing for cognitive WSNs under Wi-Fi interference. In Proceedings of the Wireless World Research Forum—WWRF 21, Stockholm, Sweden, 13–15 October 2008.
19. Ye, Z.; Memik, G.; Grosspietsch, J. Energy Detection Using Estimated Noise Variance for Spectrum Sensing in Cognitive Radio Networks. In Proceedings of the Wireless Communications and Networking Conference, Las Vegas, NV, USA, 31 March–3 April 2008; pp. 711–716.
20. Shellhammer, S.J.; Tandra, R.; Tomcik, J. Performance of power detector sensors of DVT signals in IEEE 802.22 WRANs. In Proceedings of the First International Workshop on Technology and Policy for Accessing Spectrum, Boston, MA, USA, 5 August 2006.
21. Bertsekas, D.P.; Tsitsiklis, J.N. *Introduction to Probability*; Athena Scientific: Nashua, NH, USA, 2008.
22. Peh, E.C.; Liang, Y.-C.; Guan, Y.-L.; Zeng, Y.-H. Optimization of cooperative sensing in cognitive radio networks: A sensing-throughput tradeoff view. *IEEE Trans. Veh. Technol.* **2007**, *58*, 5294–5299. [[CrossRef](#)]
23. Kuang, J. *Applied Inequalities*, 3rd ed.; Shandong Science and Technology Press: Jinan, China, 2004. (In Chinese)
24. Shevtsova, G.I. Sharpening of the upper bound of the absolute constant in the Berry-Esseen inequality. *Theory Probab. Appl.* **2007**, *51*, 549–553. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).