Energy Efficient Resource Allocation for Cognitive Radios: A Generalized Sensing Analysis

AbdulRahman Alabbasi, *Student Member, IEEE*, Zouheir Rezki, *Senior Member, IEEE*, Basem Shihada, *Senior Member, IEEE,*

Computer, Electrical and Mathematical Science and Engineering (CEMSE),

King Abdullah University of Science and Technology (KAUST),

Email: {abdulrahman.alabbasi, zouheir.rezki, basem.shihada}@kaust.edu.sa

Abstract

In this paper, we propose two resource allocation schemes for energy efficient cognitive radio systems. Our design considers resource allocation approaches that adopt spectrum sharing combined with soft-sensing information, adaptive sensing thresholds, and adaptive power to achieve an energy efficient system. We consider an energy per good-bit metric as an energy efficient objective function. Our formulation targets a multi-carrier system, such as, orthogonal frequency division multiplexing. The two resource allocation schemes, using different approaches, are designated as sub-optimal and optimal. The sub-optimal approach is attained by optimizing over a channel inversion power policy. The optimal approach utilizes the calculus of variation theory to solve a problem of instantaneous objective function subject to average and instantaneous constraints with respect to functional optimization variables. In addition to the analytical results, selected numerical results are provided to quantify the impact of soft-sensing information and the optimal adaptive sensing threshold on the system performance.

Index Terms

Spectrum sharing, Energy efficiency, Spectrum sensing, Resource allocation,

I. INTRODUCTION

Green communication is a major contributor to the global greenness. Therefore, we note the high demand on reducing the mobile phones power consumption while achieving high throughput. It is found that the radio's power consumption reaches up to 50 percent of the mobile device's power consumption [1]. Cognitive radio (CR) technology is a potential candidate to achieve greenness communication [2]. There are many proposals to achieve greenness for CR technology under both MIMO and SISO environment [3], [4], [5], [6], [7]. Recall that CR has been proposed to overcome the inefficient use of frequency spectrum and its resulting scarcity. In CR, secondary user (SU) may share the bands of the primary user (PU) through various schemes [8], [9]. Combining the knowledge of PU's existence into the SU's transmission scheme results in an improved performance. Therefore, existing work utilizes the PU's soft-sensing information in single-input single-output systems to maximize the system capacity [10], [11].

It is well known that the transmission power is one of the main factors in deciding the system communication rate, reliability, and power consumption [12]. Therefore, researchers have been working toward achieving an optimal energy efficient system, i.e., minimizing the system energy while preserving its quality of service (QoS) parameters, in terms of rate, delay, etc. There are several approaches to handle the problem by optimizing either the rate or the overall energy. Some researchers consider optimizing the total transmission power consumption while forcing the rate constraint to be above a certain threshold [13], while others consider maximizing the rate under a total power budget [14], [15]. The authors of [6] considered a downlink, multi-casting, SU system under MIMO environment. The transmission power subject to primary interference constraint is minimized as a quadratic constraint quadratic problem under imperfect channel state information (CSI). Two randomization schemes have been proposed since the problem is not a convex one. It is argued that minimizing the energy per goodbit (EPG) outperforms both categories in-terms of system energy efficiency [16], [17]. The EPG function describes the energy consumption per successfully received bit. The authors of [18] considered a throughput per energy as an energy efficiency metric. They proposed a joint design of the optimal sensing duration and transmission duration, without an optimal power allocation, to maximize their system efficiency. They guaranteed the protection of PU via limiting the detection probability within a certain threshold.

In this paper, we design and analyze an energy efficient spectrum sharing system by utilizing the sensing information about PU. Two resource allocation schemes have been proposed to show the impact of adaptive optimal power, adaptive sensing threshold, and soft-sensing information on the system performances. A description of our settings is as follows,

- We assume that the SU network is a delay intolerant network, such as, Voice over IP, Video over IP, and the legacy phone networks. On the other hand, the PU network, which appears in our problem through the average constraint, is assumed to be delay tolerant network, such as, file sharing network, ftp, and http.
- We consider the existence of a temporal database that describes the presence of PU for certain periods of time. The database acts as a source for the prior hypotheses of PU existence, not as a sensing source (it is accessed by SU in a daily or weekly basis). This database has been developed through learning algorithms or experiments and measurements which have been done by a spectrum organization via exploiting a cognitive radio engine [19], [20].
- Under the proposed system, we assume the availability of a spectrum sensor, at the SU, which collects real time information about PU signal. This assumption allows us to enforce and adjust the prior probability of PU existence which was obtained from the database. Furthermore, this sensor improves the accuracy of the probability of PU existence at the database. Most importantly, the sensing metric, resulted from the SU sensor, gives a relative measurement on how much a SU can aggregate on the corresponding sub-channel. Large SU power aggregation is allowed when the sensing metric indicates that the corresponding long-term effect on PU is small. Whereas, power aggregation is not allowed when the sensing metric indicates a high long-term effect on PU.
- We assume that the SU sensor adopt an energy detection scheme for attaining the soft-sensing information. This choice is justified by the fact that it is hard to obtain PU CSI beforehand.

- In the benchmark system, we assume no existence of an on-board spectrum sensor. This assumption highlight the gain of utilizing the adaptive sensing threshold and the soft-sensing information.
- We develop three more benchmark systems, in Section IV, to further evaluate the performance of the proposed systems.
- We consider both the EPG cumulative density function (CDF) and the outage metrics (formally defined later) to evaluate our system performance.

The proposed approaches are designated as sub-optimal and optimal approaches. The sub-optimal one is achieved via optimizing the EPG metric using a channel inversion power policy (thus noted as sub-optimal). Whereas, the optimal one utilizes the calculus of variation to minimize an instantaneous EPG subject to average and instantaneous constraints (thus noted as optimal). In our analysis, we prove the pseudo-convexity structure of the objective function (EPG) and the quasi-convexity of the corresponding constraints. This facilitate the finding of a global optimal solution. Our numerical results confirm that the optimal approach achieves better performance compared to the sub-optimal one. Analytical results are provided for both the proposed and benchmark systems. Numerical results show the improvement of the proposed schemes compared to the benchmark ones.

Unlike [16], which minimizes the EPG metric without considering the CR environment, our work protects the PU and utilizes the sensing information in order to minimize the EPG metric. In addition to the difference in the system model assumptions, the utilization of quasi-convexity and pseudo-convexity analysis of the targeted problem is distinct from that in [21], [22].

The organization of the paper is as follows. Section II describes our system model with a related background on PU sensing approach. Section III discusses the problem formulation and analyzes of both the benchmark (optimal and sub-optimal schemes) and proposed (optimal and sub-optimal schemes) systems. Finally, selected numerical results are presented in section IV.

II. System Model and Related Background

A. System Model

In this system, we consider a CR multi-carrier system where all terminals have a single antenna, i.e., single-input single-output (SISO) system. Figure 1 shows our system model for sub-channel i, where i is the index of a sub-carrier in the orthogonal frequency-division mutiplexing (OFDM) system, $i \in \{1, ..., N\}$. In Fig. 1 the fading channels between primary transmitter (PT) and secondary receiver (SR), primary receiver (PR) and secondary transmitter (ST), ST and SR are designated by h_{psi}, h_{spi}, h_{si}, and their corresponding squared modulus $\gamma_{psi} = |h_{psi}|^2$, $\gamma_{spi} = |h_{spi}|^2$, $\gamma_{si} = |h_{si}|^2$, respectively. The previously mentioned channels gains are assumed to be independent. The corresponding channel vector is expressed as $\boldsymbol{\gamma}_{ps} = \{\gamma_{ps1}, \gamma_{ps2}, \ldots, \gamma_{psN}\}$, $\boldsymbol{\gamma}_{sp} = \{\gamma_{sp1}, \gamma_{sp2}, \ldots, \gamma_{spN}\}$, and $\boldsymbol{\gamma}_{s} = \{\gamma_{s1}, \gamma_{s2}, \ldots, \gamma_{sN}\}$.

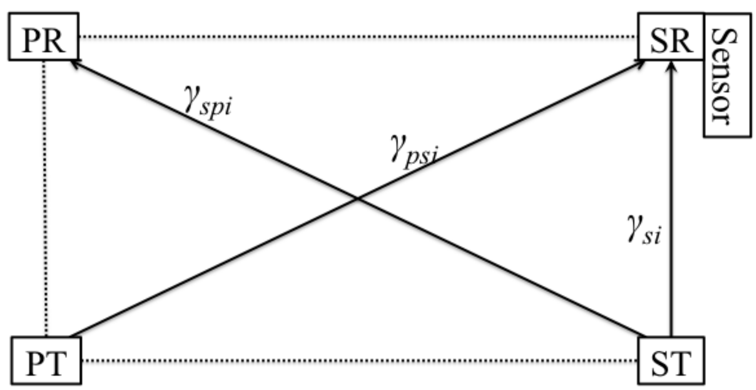

Fig. 1. System Model.

It is assumed that the sensor is on the SR side. Additionally, we consider that both ST and SR share the knowledge about the sensing information (ζ_i, defined later), the SU channel γ_{si}, and the average value of γ_{spi} ($\overline{\gamma_{spi}}$) through a feedback channel. It is assumed that PT transmits with a fixed P_p. This is justified by the fact that PT may not know the instantaneous value of its channel, and thus it is using a constant power policy. On the other hand, ST has an adaptive power allocation policy (to be described later) denoted as P_{si}, per sub-carrier. SU decodes the PU interference as noise, since it does not know γ_{psi}. Hence, it is unable to perform any interference mitigation technique. Considering that our system adopt a CR environment, we have to protect the

PU from the SU interference. Therefore, we enforce an average constraint on the ST interference toward the PR.

B. Sensing Background

In this section, we provide the necessary background for incorporating the sensing information in our system. The sensing information is associated with our optimization problem through the average constraint on the interference from ST to PR. Figure 2 shows the sensing framework, where the radio senses the channels for a period of time T_s, then transmits for a period of time T_c. In this work, we assume that $T_c >> T_s$, such that T_s does not affect our performance. Furthermore, we assume that the channel coherence time is large enough so that it does not change within two periods of sensing. In our system, we use an energy detection scheme,

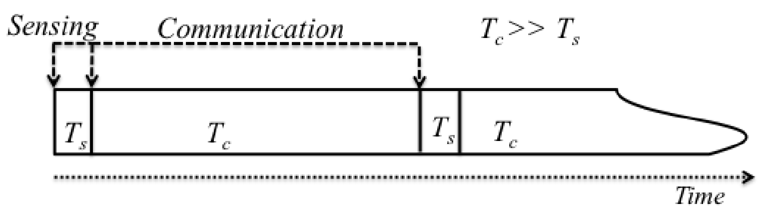

Fig. 2. Sensing Framework.

because we do not have a prior knowledge about the PU CSI. As known in the literature [23], the output of the energy detector is expressed as follows,

$$\zeta_i(N_s) = \frac{1}{N_s}\sum_{n=1}^{N_s} |y_i(n)|^2 \ . \tag{1}$$

The vector expression of the sensing information is $\boldsymbol{\zeta} = \{\zeta_1, \zeta_2, \ldots, \zeta_N\}$, and N_s is the number of sensing samples. The received signal at SR sensor, $y_i(n)$, corresponding to sub-carrier i, is expressed as follows,

$$y_i(n) = \begin{cases} \eta_i(n) & \mathcal{H}_0: \quad \text{(PU is idle)} \\ h_{psi}(n)x_i(n) + \eta_i(n) & \mathcal{H}_1: \quad \text{(PU is active)} \end{cases} , \tag{2}$$

where, $x_i(n)$ is the transmitted symbol from the PT at sub-carrier i, $\eta_i(n)$ is the Additive White Gaussian Noise (AWGN) at the sensor side at sub-carrier i.

The authors of [23] have done a thorough investigation on the probability density function (PDF) of $\zeta_i(N_s)$ for both cases $\mathcal{H}_0$ and $\mathcal{H}_1$. Following the same general approach, we obtain the PDF of both $\mathcal{H}_0$ and $\mathcal{H}_1$, which we designate as $h_0(\zeta_i)$ and $h_1(\zeta_i)$, and their corresponding means μ_0, μ_1, and variances δ_0^2, δ_1^2, respectively. Deciding whether PU is active or idle requires an optimal decision mechanism. We use the likelihood function of the PDFs of the two hypotheses as follows,

$$\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui} \ , \tag{3}$$

where, γ_{ui} is the sensing threshold which identifies the absence of PU if (3) is satisfied. By solving the inequality in (3), we obtain a set of ζ_i that lies within a region Z_0 and satisfies (3), meaning that PU is idle. On the other hand, if (3) is not satisfied, then the set of the resulting values of ζ_i lie within a region Z_1, where $Z_1 = \overline{Z_0}$, and the PU is declared to be active, so we refrain from transmitting.

III. Problem Formulation and Systems' Analysis.

This section is divided into two sub-sections. The first one discusses the sub-optimal analysis of both the benchmark and proposed systems, whereas, the second sub-section considers the optimal analysis of both systems. Before proceeding with the mentioned sub-sections, we illustrate a common problem formulation that is shared by all systems and schemes, i.e., benchmark and proposed under both sub-optimal and optimal approaches.

The EPG metric is a common objective function among all schemes. Furthermore, there are three common constraints, namely, peak power constraint, minimum rate constraint, and average PU interference constraint (CR constraint). The formulation of the common problem, called $\mathcal{P}_0$, is expressed as follows,

$$\mathcal{P}_0 : \min_{\mathbf{P_s}(\boldsymbol{\gamma})} \quad \mathcal{E}(\mathbf{P_s}(\boldsymbol{\gamma})) = \frac{k_t \sum_{i=1}^{N} P_{si}(\boldsymbol{\gamma}) + k_c}{\sum_{i=1}^{N} \log(1 + \frac{P_{si}(\boldsymbol{\gamma})\gamma_{si}}{1+P_p\gamma_{psi}})} \tag{4a}$$

$$\text{s.t.} \quad P_{si}(\boldsymbol{\gamma}) \leq P_{max}, \quad \forall i \in \{1, \ldots, N\} \tag{4b}$$

$$\sum_{i=1}^{N} \log(1 + \frac{P_{si}(\boldsymbol{\gamma})\gamma_{si}}{1 + P_p\gamma_{psi}}) \geq R_{min} \tag{4c}$$

$$\mathbb{E}_{\gamma_i}\left[P_{si}\left(\boldsymbol{\gamma}\right)\gamma_{spi} - P_{ave}\right] \leq 0, \quad \forall i \in \{1,\ldots,N\}\ . \tag{4d}$$

The EPG objective function is defined as the power to rate ratio, as in (4a). The power profile $P_{si}\left(\boldsymbol{\gamma}\right)$ and P_p are the transmission powers of SU and PU, respectively, at sub-carrier i, N is the total number of sub-carriers. The power profile $P_{si}\left(\boldsymbol{\gamma}\right)$ is defined individually for the benchmark and proposed systems in each of the following sub-sections. The corresponding vector of the power profile is expressed as $\mathbf{P_s}\left(\boldsymbol{\gamma}\right) = \{P_{s1}\left(\boldsymbol{\gamma}\right), P_{s2}\left(\boldsymbol{\gamma}\right), \ldots, P_{sN}\left(\boldsymbol{\gamma}\right)\}$. The parameters P_{max} and R_{min} are assigned constants for the SU's instantaneous power budget per sub-carrier and minimum fixed rate of the SU, respectively. The constants k_t and k_c refers to the power amplifier and circuit operation constant powers, respectively. The variable γ_i is defined separately in the proposed systems and the benchmark ones. In the proposed systems γ_i consists of γ_{si}, γ_{spi}, and the sensing metric variable ζ_i, whereas, under the benchmark systems γ_i consists only of γ_{si} and γ_{spi} only. The vector notation of γ_i is expressed as $\boldsymbol{\gamma} = \{\gamma_1, \ldots, \gamma_N\}$. P_{ave} is the average power constraint on the received interference by PR from ST, and it is defined as,

$$P_{ave} = \begin{cases} Q_{int}, & \frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui} \bigcap \text{PU is ON} \\ P, & \frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui} \bigcap \text{PU is OFF} \end{cases}. \tag{5}$$

Recall that $\Pr\{\text{PU is ON}\}$ and $\Pr\{\text{PU is OFF}\}$ are the probabilities of PU being active and idle, respectively, obtained prior from the database. The objective is to tune the interference threshold P_{ave} depending on the sensing information and the prior information. Intuitively, if PU is ON, then P_{ave} is set to the lowest value, otherwise, P_{ave} can be arbitrary high. It is then clear that $Q_{int} \leq P$. In case we do not have sensing information about PU existence (as in the benchmark cases) we consider the worst case (to protect the PR) and use $P_{ave} = Q_{int}$. It is also worth mentioning that in case $\frac{h_0(\zeta_i)}{h_1(\zeta_i)} < \gamma_{ui}$ we stop transmitting, i.e., $P_{si} = 0$, therefore, in this case it is not necessary to enforce an interference constraint. Constraint (4b) is the peak power constraint over each sub-carriers, constraint (4c) is the minimum required SU rate constraint, and (4d) is the average power and interference constraint. This formulation of an instantaneous objective function with an average interference constraint is justified by the earlier assumption. Where the PU system operate under a delay tolerant constraints, whereas, the SU is intolerant toward the delay.

Note that the structure of problem (4) is complicated to solve for several reasons as follows. The objective function (4a) is an instantaneous function, whereas constraint (4d) is an average constraint. This is a difficult problem if the term inside the optimization variables are a functional of functions, $\mathbf{P_s}(\gamma_s)$. In Sec. III-B we solve this problem through a series of steps and using the calculus of variation theory. Furthermore, the objective function (4a) is clearly a fractional non-convex function on P_{si}. Therefore, it is essential to show a technique that provide a global optimal solution to the problem. It is known that by satisfying the Karush-Kuhn-Tucker (KKT) conditions a global optimal solution can be obtained for a convex problem, i.e., convex objective function with convex constraint. A generalized version of this result is stated in the theorem below [24][Sec. 4.4]. Prior to Theorem 1 it is necessary to define the following,

Definition Problem $\mathcal{P}$ is defined as follows,

$$\mathcal{P}:\ \min_{x} f(x), x \in S = \left\{ x \in X : g_i(x) \leq 0, i \in \{1, \dots, m\} \right\}. \tag{6}$$

Theorem 1. *Consider problem $\mathcal{P}$ with x_0 as a feasible point. Let f be pseudo-convex at $x_0 \in S$ and that g_i are differentiable and quasi-convex at x_0. If there exist $\lambda_i \in \mathbb{R}$, such that,*

$$\begin{cases} \nabla f(x_0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x_0) = 0,\ i \in \{1, \dots, m\} \\ \lambda_i \geq 0,\ i \in \{1, \dots, m\} \\ \lambda_i g_i(x_0) = 0,\ i \in \{1, \dots, m\} \end{cases}, \tag{7}$$

then x_0 is a global minimum point for $\mathcal{P}$.

Proof: The proof is stated in [24][Sec. 4.4]. ■

Theorem 1 is utilized to solve problem 4. The following lemma is deducted to show that problem 4 satisfies the pseudo-convexity and quasi-convexity settings stated in Theorem 1.

Lemma 1. *Let $\mathbf{P}_\mathbf{s}^*(\gamma)$ be an optimal solution, of problem $\mathcal{P}_0$, that satisfies the KKT conditions. It follows that $\mathbf{P}_\mathbf{s}^*(\gamma)$ is a global optimal solution for problem $\mathcal{P}_0$.*

Proof: The proof is in Appendix A ■

A. Sub-Optimal Approach: Channel Inversion

In this section, we address the analysis of the benchmark system through the sub-optimal approach. Then, we proceed with the analysis of the proposed system through the sub-optimal approach.

1) Benchmark System Analysis: In this section, we analyze the benchmark system, which does not consider the sensing information, using the sub-optimal approach. The benchmark problem formulation is similar to the one described in problem $\mathcal{P}_0$, where (4d) is now expressed as,

$$\mathbb{E}_{\gamma_{si},\gamma_{spi}}\{P_{si}\gamma_{spi}\} \leq Q_{int} \ . \tag{8}$$

In the benchmark case, we do not have sensing information about PU existence, hence, we consider the worst case (to protect the PR) scenario where we let $P_{ave} = Q_{int}$. Let us call the new problem after changing (4d) into (8) as $\mathcal{P}_1$, formulated as follows,

$$\mathcal{P}_1 : \min_{\mathbf{P_s}(\boldsymbol{\gamma})} \qquad (4a) \tag{9a}$$

$$\text{s.t.} \qquad (4b);\ (4c);\ (8)\ . \tag{9b}$$

Now, we focus on the channel inversion power policy, which is defined for the benchmark system as follows,

$$P_{si}\left(\boldsymbol{\gamma}\right) = \begin{cases} 0, & \gamma_{si} < \gamma_{vi} \\ \frac{\sigma_i}{\gamma_{si}}, & \gamma_{si} \geq \gamma_{vi} \end{cases} , \tag{10}$$

where γ_{vi} is a SU's channel quality threshold (to be optimized) and σ_i is the power inversion coefficient (to be optimized). The corresponding vector notation is $\boldsymbol{\sigma} = \{\sigma_1, \ldots, \sigma_N\}$. The variable γ_{vi} gives an indication about the SU channel γ_{si}; whether it is in deep fade, so we stop transmission, or in good condition, so we continue transmission. This power policy is defined in this form to preserve a fair comparison to the proposed scheme in III-A2.

In order to solve problem $\mathcal{P}_1$, we begin by solving the expectation in (8). Since both γ_{si} and γ_{spi} are independent variables, while taking into account the assumed power policy in (10), we can separate both integrals as follows,

$$\sigma_i \int_{\gamma_{vi}}^{\infty} \frac{1}{\gamma_{si}} f_{\gamma_{si}}(\gamma_{si}) d\gamma_{si} \int_{0}^{\infty} \gamma_{spi} f_{\gamma_{spi}}(\gamma_{spi}) d\gamma_{spi} \leq Q_{int} \ . \tag{11}$$

The first integral results in a function of γ_{vi}, $G(\gamma_{vi}) = \int_{\gamma_{vi}}^{\infty} \frac{1}{\gamma_{si}} f_{\gamma_{si}}(\gamma_{si}) d\gamma_{si}$, which depends on the PDF of γ_{si}. The second integral results in $\overline{\gamma_{spi}}$, the average value of the γ_{spi}. Thus, (11) might be written as,

$$\overline{\gamma_{spi}}\, \sigma_i \, G(\gamma_{vi}) \leq Q_{int} \ . \tag{12}$$

In order to utilize Theorem 1, in obtaining a global optimal solution to the benchmark problem, it is necessary to verify that constraint (12) is jointly quasi-convex with respect to both $\boldsymbol{\sigma}$ and $\boldsymbol{\gamma_v}$. This can be verified via two steps. The first step is to show that the function $G(\gamma_{vi})$ is a non-increasing function with respect to γ_{vi}. This is deducted by using Leibniz rule on $G(\gamma_{vi})$, i.e., $\frac{\partial}{\partial \gamma_{vi}} G(\gamma_{vi}) = \frac{\partial}{\partial \gamma_{vi}} \int_{\gamma_{vi}}^{\infty} \frac{1}{\gamma_{si}} f_{\gamma_{si}}(\gamma_{si}) d\gamma_{si} = -\frac{1}{\gamma_{vi}} f_{\gamma_{vi}}(\gamma_{vi})$. In the second step, we show that the product of a non-increasing function and a quasi-convex function results in a quasi-convex function, as in the following lemma. Note that, in the following lemma, the notation of all optimization variables is replaced by a vector $\mathbf{x}$, for notation generality.

Lemma 2. *Let $f(\mathbf{x})$ be a quasi-convex function and $g(\mathbf{x})$ be a non-increasing function. Define a function $z(\mathbf{x})$ as the product between $f(\mathbf{x})$ and $g(\mathbf{x})$, such that, $z(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$. Then, $z(\mathbf{x})$ is a quasi-convex function with respect to $\mathbf{x}$.*

Proof: The proof is given in Appendix B ■

Knowing that $G(\gamma_{vi})$ is a non-increasing function and $\overline{\gamma_{spi}}\, \sigma_i$ is a quasi-convex function, hence, it is verified that constraint (12) is jointly quasi-convex. Thus, satisfying the KKT conditions is enough to obtain a global optimal solution to the corresponding problem.

The benchmark problem, $\mathcal{P}_1$, is reformulated by introducing a new optimizing variable t. Let us note the new problem, after introducing t, as $\mathcal{P}'_1$. Hereafter, we use the definition of equivalency between two problems as in [25], (i.e., "Two problems are equivalent if from a solution of one, a solution of the other is readily found, and vice versa”). Problem $\mathcal{P'}_1$ is expressed as follows,

$$\mathcal{P'}_1 : \min_{t,\boldsymbol{\sigma},\boldsymbol{\gamma}_v} \quad t \tag{13a}$$

$$\text{s.t.} \frac{k_t \sum_{i=1}^{N} P_{si}\left(\boldsymbol{\gamma}\right) + k_c}{\sum_{i=1}^{N} \log\left(1 + \frac{P_{si}(\boldsymbol{\gamma})\gamma_{si}}{1+P_p\gamma_{psi}}\right)} \leq t \tag{13b}$$

$$P_{si} \leq P_{max}, \quad \forall i \in \{1,\ldots,N\} \tag{13c}$$

$$\sum_{i=1}^{N} \log(1 + \frac{P_{si}\gamma_{si}}{1+P_p\gamma_{psi}}) \geq R_{min} \tag{13d}$$

$$\overline{\gamma_{spi}}\sigma_i G(\gamma_{vi}) \leq Q_{int}, \quad \forall i \in \{1,\ldots,N\} \tag{13e}$$

$$\gamma_{vi} \leq \gamma_{si} \ . \tag{13f}$$

Note that constraints (13c), (13d), (13e), and (13f) are equivalent to constraints (4b), (4c), (8) and the inequality condition in (10), respectively.

Proposition 1. *Problem $\mathcal{P}'_1$ is equivalent to problem $\mathcal{P}_1$.*

Proof: The proof is presented in Appendix C. ■

Note that problem $\mathcal{P}'_1$ is also called the epigraph form of $\mathcal{P}_1$. Therefore, by definition, problems $\mathcal{P}'_1$ and $\mathcal{P}_1$ are equivalent. The global optimal solution is found by utilizing the Lagrangian multiplier method, i.e., constructing the Lagrangian function and satisfying the KKT conditions. The solution of problem (13) is stated in the following theorem.

Theorem 2. *The optimal values of the three optimization variables of problem $\mathcal{P'}_1$ are obtained as follows,*

- *The optimal value of σ_i has two distinct expressions depending on the value of λ_{2i}, which is the Lagrangian multiplier of constraint (13e), and they are obtained as follows,*
  - $\lambda_{2i} = 0 \implies \frac{\sigma_i^*}{\gamma_{si}} = \min\left(\left[\frac{\lambda_1 t + \lambda_3}{k_t\lambda_1} - \frac{1+P_p\gamma_{psi}}{\gamma_{si}}\right]^+, P_{max}\right)$, *where λ_1 and λ_3 are the Lagrangian multipliers of constraint (13b) and (13d), respectively.*
  - $\lambda_{2i} \neq 0 \implies \sigma_i^*$ *is expressed as the solution of a quadratic equation as,* $\sigma_i^* = \min\left(\left[\frac{-b\pm\sqrt{(b^2-4ac)}}{2a}\right]^+, \sigma_{mn}\right)$, *where,* $a = \frac{1}{\gamma_{si}}$, $b = \left[\frac{1+\gamma_{psi}P_p}{\gamma_{si}} + \frac{\lambda_{2i}}{\lambda_1}Q_{int} - \frac{(\lambda_1 t+\lambda_3)}{k_t\lambda_1}\right]$,

and, $c = \left[\frac{\lambda_{2i}}{k_t \lambda_1} Q_{int}(1+\gamma_{psi}P_p)\right]$, $\sigma_{mn} = \min(\sigma_n^+, P_{max}\gamma_{si})$, *and,* $\sigma_n^+ = \frac{Q_{int}}{G(\gamma_{vi})\overline{\gamma_{spi}}}$.

- *The optimal value of* γ_{vi} *is found by solving the following differential equation, which depends on* $G(\gamma_{vi})$*,*

$$\frac{\partial}{\partial \gamma_{vi}} G(\gamma_{vi}) = -\frac{\lambda_{4i}}{\lambda_{2i}\overline{\gamma_{spi}}\sigma_i} \ , \tag{14}$$

where λ_{4i} *is the Lagrangian multiplier associated with constraint (13f)*

- *The optimal value of* t *is found numerically for general* N *by solving the following equation,*

$$\prod_{i=1}^{N}\left(1+\frac{\sigma_i^*}{1+P_p\gamma_{psi}}\right) - \exp(1/\lambda_1) = 0 \ . \tag{15}$$

Proof: The proof is presented in Appendix D. ■

The analytical results of Theorem 2 for the optimization variable σ_i takes either a water-filling shape ($\lambda_{2i} = 0$) or a quadratic form (if $\lambda_{2i} \neq 0$). It is observed from (14) that the SU channel quality threshold γ_{vi} depends on how critical both the constraints (13e) and (13f) are, which affect the values of λ_{2i} and λ_{4i}, respectively. Furthermore, the value of γ_{vi} depends on the transmit power variable (σ_i). The Lagrangian multipliers $\lambda_1, \lambda_{2i}, \lambda_3, \lambda_{4i}$ are obtained via numerically solving the corresponding KKT conditions.

2) Proposed System Analysis: In this section, we formulate and analyze the proposed problem, using the sub-optimal approach. In order to show the sensing effect, we consider the general case where PU is OFF (idle) and PU is ON (active), while considering the two sensing hypotheses, $\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \lessgtr_{\mathcal{H}_0}^{\mathcal{H}_1} \gamma_{ui}$. The new problem follows a similar formulation as problem $\mathcal{P}_0$, where constraint (4d) is now expressed as,

$$\mathbb{E}_{\gamma_{si},\gamma_{spi},\zeta_i}\{P_{si}(\boldsymbol{\gamma})\,\gamma_{spi} - P_{ave}\} \leq 0 \ . \tag{16}$$

The parameter P_{ave} is included in the expectation because the choice of P_{ave} is based on the sensing metric ζ_i, recall (5). We call the new problem that substitutes (4d) in $\mathcal{P}_0$ by (16) as $\mathcal{P}_2$. Our optimization variables of $\mathcal{P}_2$ are: SU transmission power ($P_{si}(\boldsymbol{\gamma})$), sensing metric threshold (γ_{ui}), and SU channel quality threshold (γ_{vi}). Since P_{si} is a function of three random variables, $\boldsymbol{\gamma}_s, \boldsymbol{\gamma}_{sp}$, and $\boldsymbol{\zeta}$, and the expectation in (16) is a function of P_{si}, it is clear that we are dealing with a functional of functions. Therefore, for simplicity in solving inequality (16), we utilize the channel inversion power policy as follows,

$$P_{si} = \begin{cases} 0, & \frac{h_0(\zeta_i)}{h_1(\zeta_i)} < \gamma_{ui} \ \cup \ \gamma_{si} < \gamma_{vi} \\ \frac{\sigma_i}{\gamma_{si}}, & \frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui} \ \cap \ \gamma_{si} \geq \gamma_{vi} \end{cases}. \tag{17}$$

It is seen that the power policy in (17) reduces to that of the benchmark system in (10) if the sensing constraint $\left(\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui}\right)$ is ignored. In order to solve $\mathcal{P}_2$, we reformulate (16) into a convenient form, as summarized in lemma 3.

Lemma 3. *Considering the power policy mentioned in (17) and the three averaging variables γ_{si}, γ_{spi}, and ζ_i, constraint (16) is reformulated as follows,*

$$\sigma_i \overline{\gamma_{spi}} G(\gamma_{vi}) \left[\alpha K_{01}(\gamma_{ui}) + \overline{\alpha} K_{00}(\gamma_{ui})\right] \leq \left[\alpha K_{01}(\gamma_{ui}) Q_{int} + \overline{\alpha} K_{00}(\gamma_{ui}) P\right], \tag{18}$$

where $K_{01}(\gamma_{ui}) = \int_{Z_0} h_1(\zeta_i) d\zeta_i$, $K_{00}(\gamma_{ui}) = \int_{Z_0} h_0(\zeta_i) d\zeta_i$.

Proof: Utilizing (17) the first term of constraint (16) is rewritten as,

$$\begin{aligned} \mathbb{E}_{\gamma_{si},\gamma_{spi},\zeta_i}\left[\frac{\sigma_i}{\gamma_{si}}\gamma_{spi}\right] &= \alpha \mathbb{E}_{\gamma_{si},\gamma_{spi},\zeta_i}\left[\frac{\sigma_i}{\gamma_{si}}\gamma_{spi} | \text{PU is ON}, \frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui}\right] \Pr\left[\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui} | \text{PU is ON}\right] \\ &+ \overline{\alpha} \mathbb{E}_{\gamma_{si},\gamma_{spi},\zeta_i}\left[\frac{\sigma_i}{\gamma_{si}}\gamma_{spi} | \text{PU is OFF}, \frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui}\right] \Pr\left[\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui} | \text{PU is OFF}\right], \end{aligned} \tag{19}$$

where, for the ease of notation, $\alpha = \Pr[\text{PU is ON}]$ and $\overline{\alpha} = \Pr[\text{PU is OFF}]$. The notation $\Pr[.]$ is the probability of the in-bracket event. We begin by solving the first term of (19), which consists of two unknown quantities and one known probability α. The second term can be solved similarly. We utilize the independence among γ_{si}, γ_{spi}, and ζ_i to solve the expectation in (19) as follows,

$$\begin{aligned} &\mathbb{E}_{\gamma_{si},\gamma_{spi},\zeta_i}\left[\frac{\sigma_i}{\gamma_{si}}\gamma_{spi} | \text{PU is ON}, \frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui}\right] = \\ &\sigma_i \int_{\gamma_{vi}}^{\infty} \frac{1}{\gamma_{si}} f_{\gamma_{si}}(\gamma_{si}) d\gamma_{si} \int_0^{\infty} \gamma_{spi} f_{\gamma_{spi}}(\gamma_{spi}) d\gamma_{spi} \int_{Z_0} h_1(\zeta_i) d\zeta_i = \sigma_i \overline{\gamma_{spi}} G(\gamma_{vi}) K_{01}(\gamma_{ui}), \end{aligned} \tag{20}$$

where, Z_0 is as defined in Sec. II-B. Both the first and the second integrals can be solved in a similar way as solving (11) to obtain (12). As for the third integral, it can be represented as a function of γ_{ui}, $K_{01}(\gamma_{ui}) = \int_{Z_0} h_1(\zeta_i) d\zeta_i$. The other two terms in (19) are expressed as follows,

$$\Pr\left[\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui}|\text{PU is ON}\right] = K_{01}(\gamma_{ui}) \quad \text{and} \quad \Pr\left[\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui}|\text{PU is OFF}\right] = K_{00}(\gamma_{ui}). \tag{21}$$

The second term of constraint (16) is rewritten by considering (5) as follows,

$$\mathbb{E}_{\gamma_{si},\gamma_{spi},\zeta_i}\left[P_{ave}\right] = \left[\alpha\Pr\left[\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui}|\text{PU is ON}\right] Q_{int} + \overline{\alpha}\Pr\left[\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \geq \gamma_{ui}|\text{PU is OFF}\right] P\right]. \tag{22}$$

Finally, (18) in Lemma 3 can be deducted by combining (19), (22), (20), and (21). ■

Comparing (18) and (12) the difference between the proposed system and the benchmark system is observed. In (18) a sensing related function appears in the P and Q_{int} terms, which allows us to aggregate more on the PU channel. It is observed that the function $K_{00}(\gamma_{ui})$ acts as the complement probability of the false alarm, i.e., correctness probability. Whereas, $K_{01}(\gamma_{ui})$ acts as a miss detection probability, i.e., errorness probability. It is clear that increasing $\frac{h_0(\zeta_i)}{h_1(\zeta_i)}$ results in increasing the ζ region Z_0. Therefore, increasing $K_{00}(\gamma_{ui})$, which increasees the feasibility region of constraint (18). On the other hand, increasing γ_{ui} results in decreasing both $K_{00}(\gamma_{ui})$ and $K_{01}(\gamma_{ui})$ which reduces the the feasibility region of constraint (18) up to a certain level depending on the values of the parameters Q_{int} and P.

In order to utilize Theorem 1 in solving problem $\mathcal{P}_2$, it is necessary to verify that constraint (18) is a jointly quasi-convex with respect to σ_i, γ_{vi}, γ_{ui}. Earlier, in Lemma 2, it is shown that, without introducing the sensing variable γ_{ui}, the original constraint is a quasi-convex one. The quasi-convexity property of constraint (18) is verified in two steps, in a similar line with the quasi-convexity proof of (12). The first step is to show that the sensing information part of the constraint, depends on γ_{ui}, is non-increasing. The second step is to show that the product of $(\sigma_i\overline{\gamma_{spi}}G(\gamma_{vi}))$ and $\left(\frac{[\alpha K_{01}(\gamma_{ui})+\overline{\alpha}K_{00}(\gamma_{ui})]}{[\alpha K_{01}(\gamma_{ui})Q_{int}+\overline{\alpha}K_{00}(\gamma_{ui})P]}\right)$ has a quasi-convex structure. It is proved, earlier, that $(\sigma_i\overline{\gamma_{spi}}G(\gamma_{vi}))$ has a quasi-convex structure. Furthermore, it is shown in Appendix E that $\left(\frac{[\alpha K_{01}(\gamma_{ui})+\overline{\alpha}K_{00}(\gamma_{ui})]}{[\alpha K_{01}(\gamma_{ui})Q_{int}+\overline{\alpha}K_{00}(\gamma_{ui})P]}\right)$ is a non-increasing function with respect to γ_{ui}. Thus, Lemma 2 verifies the quasi-convex structure of constraint (18). Hence, problem $\mathcal{P}_2$, including constraint

(18), satisfies Theorem 1. Therefore, satisfying the KKT conditions is enough to obtain the global optimal solution.

Problem $\mathcal{P}_2$ is converted to $\mathcal{P}_2'$ in a similar way to the conversion of $\mathcal{P}_1$ to $\mathcal{P}_1'$. Furthermore, we follow similar steps of Theorem 2's proof in constructing the Lagrangian function of $\mathcal{P}_2'$ and satisfying the KKT conditions. It follows that the optimal value of σ_i is derived in the following Proposition.

Proposition 2. *Considering the power policy mentioned in (17), the optimal value of* $\frac{\sigma_i}{\gamma_{si}}$*, of the proposed system formulated in* $\mathcal{P}_2$*, is derived as follows,*

$$\frac{\sigma_i^*}{\gamma_{si}} = \min\left(\left[\frac{(\lambda_1 t + \lambda_3)}{k_t\lambda_1 + \lambda_{2i} G(\gamma_{vi})\overline{\gamma_{spi}}[\alpha K_{01}(\gamma_{ui}) + \overline{\alpha} K_{00}(\gamma_{ui})]} - \frac{(1 + P_p\gamma_{psi})}{\gamma_{si}}\right]^+, P_w^+\right), \quad (23)$$

where, $P_w^+ = \min\left(\frac{\sigma_w}{\gamma_{si}}, P_{max}\right)$*, and* $\sigma_w = \frac{\alpha K_{01}(\gamma_{ui})Q_{int} + \overline{\alpha}K_{00}(\gamma_{ui})P}{G(\gamma_{vi})\overline{\gamma_{spi}}[\alpha K_{01}(\gamma_{ui}) + \overline{\alpha}K_{00}(\gamma_{ui})]}$.

Proof: The proof of Proposition 2 is done in a similar way to the one of Theorem 2, except that we use constraint (18) instead of constraint (13e) to construct our Lagrangian function. ■

The first term of the minimum function in (23) can be interpreted as in the water-filling power policy with the addition of the sensing information effect, expressed in $[\alpha K_{01}(\gamma_{ui}) + \overline{\alpha}K_{00}(\gamma_{ui})]$, and the SU's channel quality effect, expressed in $G(\gamma_{vi})$. The effect of sensing reduces the power when constraint (18) is critically satisfied. On the other hand, the allocated power is increased when the term $[\alpha K_{01}(\gamma_{ui}) + \overline{\alpha}K_{00}(\gamma_{ui})]$ is decreased; we know that this term is inversely proportional to γ_{ui}. Taking into consideration that increasing the soft-sensing information parameter, $\frac{h_0(\zeta_i)}{h_1(\zeta_i)}$, results in increasing the correctness probability and reducing the errorness probability. Thus, optimizing γ_u will be also controlled by $\frac{h_0(\zeta_i)}{h_1(\zeta_i)}$.

In order to obtain the optimal expression for the other two variables, γ_{vi} and t, we follow similar procedures as in (14) and (15), respectively. It is difficult to obtain an analytical expression for γ_{ui}, because of the structure of both functions $K_0 0\,(\gamma_{ui})$ and $K_0 1\,(\gamma_{ui})$. Therefore, the global optimal value of γ_{ui} is obtained through a numerical method, such as, bisectional method. This is easily done, and guaranteed, since constraint (18), which contains the non-convex complicated function of γ_{ui}, is a quasi-convex constraint, as proven earlier.

B. Optimal Approach

In this section, we begin by introducing a common system analysis, through an optimal approach, for both the benchmark and the proposed systems. Then, we tackle the analysis of the benchmark system. Finally, we proceed with the analysis of the proposed system.

1) Common Problem Formulation & Analysis under an Optimal Approach: In Sec. III-A2, we utilized the sub-optimal approach in order to avoid dealing with an optimization problem that involves a functional of functions variables under instantaneous objective function with an average constraint. However, in this section, we directly tackle this problem through the calculus of variation principle. A necessary background on the calculus of variation is explained in Appendix F. We begin our common problem analysis by manipulating our problem, $\mathcal{P}_0$, to be suitable for the usage of Euler-Lagrange theorem (Appendix F). The main difference between our problem and the framework of Euler-Lagrange formula is the structure of the objective function and some constraints. Our problem, $\mathcal{P}_0$, has an expectation (integration) in the averaged interference constraint (4d), thus, we need to introduce the integral in the objective function. Considering a similar approach to the one used in converting $\mathcal{P}_1$ into $\mathcal{P}_1'$, our objective function is equivalent to

$$\min_{t,\mathbf{P_s}(\boldsymbol{\gamma})} \quad t \tag{24a}$$

$$\text{s.t.} \quad k_t \sum_{i=1}^{N} P_{si}\left(\boldsymbol{\gamma}\right) + k_c - t\left(\sum_{i=1}^{N} \log\left(1 + \frac{P_{si}\left(\boldsymbol{\gamma}\right)\gamma_{si}}{1 + P_p \gamma_{psi}}\right)\right) \leq 0, \tag{24b}$$

with (24b) as additional constraint. Similar to Sec. III-A2 and Sec. III-A1 the variable $\boldsymbol{\gamma}$ is substituted by $\boldsymbol{\gamma}_s$ and $\boldsymbol{\gamma}_{sp}$ in the benchmark system case, and it is substituted by $\boldsymbol{\gamma}_s$, $\boldsymbol{\gamma}_{sp}$, and $\boldsymbol{\zeta}$ in the proposed system case.

We incorporate the integration to the objective function. Fortunately, our new objective function (24a) is not a function of the averaging random variables γ_{si}, γ_{spi}, and ζ_i, thus, the expectation of t w.r.t. them is still t. After some manipulations based on Appendix F the original problem becomes,

$$\mathcal{P}_3: \quad \min_{t,\mathbf{P_s}(\boldsymbol{\gamma})} \quad \mathbb{E}_{\gamma_i}\left[t\right] \tag{25a}$$

$$\text{s.t.} \quad \mathbb{E}_{\gamma_i}\left[v_0^2 + k_t \sum_{i=1}^{N} P_{si}(\boldsymbol{\gamma}) + k_c - t\left(\sum_{i=1}^{N} \log\left(1 + \frac{P_{si}(\boldsymbol{\gamma})\gamma_{si}}{1 + P_p\gamma_{psi}}\right)\right)\right] = 0 \quad (25b)$$

$$\mathbb{E}_{\gamma_i}\left[v_2^2 - \sum_{i=1}^{N} \log\left(1 + \frac{P_{si}(\boldsymbol{\gamma})\gamma_{si}}{1 + P_p\gamma_{psi}}\right) + R_{min}\right] = 0 \quad (25c)$$

$$\mathbb{E}_{\gamma_i}\left[P_{si}(\boldsymbol{\gamma})\gamma_{spi} - P_{ave}\right] \leq 0 \quad \forall i \in \{1, \ldots, N\}, \quad (25d)$$

where, v_0^2 and v_2^2 are dummy variables defined in Appendix F. Problem $\mathcal{P}_3$ is easily solved through the Lagrangian method, in a similar way as $\mathcal{P}_1'$ and $\mathcal{P}_2'$.

2) Benchmark System Analysis: This section shows the derivation of the optimal resource allocation of the benchmark system under the optimal approach. The optimization problem of the benchmark case is expressed as in (25), where we substitute γ_i by both γ_{si} and γ_{spi}. The construction of the corresponding Lagrangian function is done in a similar way to the previously solved problems. The expression of the optimal $P_{si}(\boldsymbol{\gamma})$ and t is shown in Proposition 3 and (28), respectively.

Proposition 3. *The optimal value of the allocated power $P_{si}(\boldsymbol{\gamma})$ for the benchmark system under the optimal approach is derived as follows,*

$$P_{si}^*(\boldsymbol{\gamma}) = \min\left(\left[\frac{(\lambda_1 t + \lambda_3)}{(k_t\lambda_1 + \lambda_2\overline{\gamma_{spi}})} - \frac{(1 + P_p\gamma_{psi})}{\gamma_{si}}\right]^+, P_n^+\right), \quad (26)$$

where λ_1, λ_2, and λ_3 are the Lagrangian multipliers corresponding to (25b), (25d), and (25c), respectively. $P_n^+ = \min(P_{ni}, P_{max})$, $P_{ni} = \frac{Q_{int}}{\gamma_{spi}}$

Proof: Following similar approach of generating the Lagrangian function in Theorem 2 and utilizing the calculus of variation techniques in Appendix F, our new modified Lagrangian function is expressed as,

$$L = t + \lambda_1\left[v_0^2 + k_t\sum_{i=1}^{N} P_{si}(\gamma_i) + k_c - t\sum_{i=1}^{N}\log(1 + \frac{P_{si}(\gamma_i)\gamma_{si}}{1 + P_p\gamma_{psi}})\right] + \sum_{i=1}^{N}\lambda_{2i}\left[P_{si}(\gamma_i)\overline{\gamma_{spi}} - Q_{int}\right] + \lambda_3\left[v_2^2 - \sum_{i=1}^{N}\log(1 + \frac{P_{si}(\gamma_i)\gamma_{si}}{1 + P_p\gamma_{psi}}) + R_{min}\right]. \quad (27)$$

Following similar steps as in the previous sections, $\frac{\partial L}{\partial P_{si}} = 0$, we obtain the optimal power allocation as in (26) ■

It is seen that the solution in (26) has a water-filling solution form. However, it does not depend either on the sensing threshold (γ_{ui}) or on the SU channel quality threshold (γ_{vi}). This is due to the fact the the optimal benchmark scheme does not exploit either the sensing information or a channel inversion power policy.

The optimal value of variable t is found numerically for general N as in (15), whereas, for $N = 1$, it is found explicitly as follows,

$$t^* = \exp(1/\lambda_1) \frac{(1 + P_p\gamma_{psi})(k_t\lambda_1 + \overline{\gamma_{spi}}\lambda_2)}{\gamma_{si}\lambda_1} - \frac{\lambda_3}{\lambda_1}. \tag{28}$$

This can be shown by solving the equation $\frac{\partial L}{\partial t} = 0$, using (27). It is observed that the t in (28) does not depend directly on other optimization variables, but indirectly through the Lagrangian multipliers.

3) Proposed System Analysis: In this section we consider the analysis of the proposed system using the optimal approach. We take into account both ON (active) and OFF (idle) hypotheses of PU with the associated probability α and $\overline{\alpha}$, respectively. Furthermore, the sensing information is considered through $\frac{h_0(\zeta_i)}{h_1(\zeta_i)} \lessgtr_{\mathcal{H}_0}^{\mathcal{H}_1} \gamma_{ui}$. Our targeted problem is equivalent to $\mathcal{P}_3$ while replacing $\boldsymbol{\gamma}_i$ by three random variable γ_{si}, γ_{spi}, and ζ_i. In order to obtain the optimal solution of our problem, we analyze the average interference constraint in (25d) in a similar way to the one in (18) while utilizing the Euler-Lagrange theorem, Appendix F. The analytical solution of our problem is stated in Proposition 4, as follows.

Proposition 4. *The optimal allocated power for the proposed system, (25), using the optimal approach is derived as follows,*

$$P_{si}^* = \min\left(P_w^+, \left[\frac{(\lambda_1 t + \lambda_3)}{k_t\lambda_1 + \lambda_{2i}\overline{\gamma_{spi}}[\alpha K_{01}(\gamma_{ui}) + \overline{\alpha}K_{00}(\gamma_{ui})\frac{h_0(\zeta_i)}{h_1(\zeta_i)}]} - \frac{(1 + P_p\gamma_{psi})}{\gamma_{si}}\right]^+\right), \tag{29}$$

where, $P_w^+ = \min(P_w, P_{max})$*, and* $P_w = \frac{\left[\alpha K_{01}(\gamma_{ui})Q_{int} + \overline{\alpha}K_{00}(\gamma_{ui})\frac{h_0(\zeta_i)}{h_1(\zeta_i)}P\right]}{\overline{\gamma_{spi}}\left[\alpha K_{01}(\gamma_{ui}) + \overline{\alpha}K_{00}(\gamma_{ui})\frac{h_0(\zeta_i)}{h_1(\zeta_i)}\right]}$.

Proof: Following the same approach of generating the Lagrangian function in Theorem 2 and utilizing (19), (22) and the calculus of variation techniques in Appendix F, the new modified Lagrangian function is expressed as,

$$L = t + \lambda_1 \left[v_0^2 + k_t \sum_{i=1}^{N} P_{si} + k_c - t \sum_{i=1}^{N} \log(1 + \frac{P_{si}\gamma_{si}}{1 + P_p\gamma_{psi}}) \right]$$
$$+ \lambda_3 \left[v_2^2 - \sum_{i=1}^{N} \log(1 + \frac{P_{si}\gamma_{si}}{1 + P_p\gamma_{psi}}) + R_{min} \right]$$
$$+ \sum_{i=1}^{N} \lambda_{2i} \left[P_{si}\overline{\gamma_{spi}} \left(\alpha K_{01}(\gamma_{ui}) + \overline{\alpha} K_{00}(\gamma_{ui}) \frac{h_0(\zeta_i)}{h_1(\zeta_i)} \right) - \left(\alpha K_{01}(\gamma_{ui}) Q_{int} + \overline{\alpha} K_{00}(\gamma_{ui}) P \frac{h_0(\zeta_i)}{h_1(\zeta_i)} \right) \right]. \quad (30)$$

Similar to Sec. III-A2, it can be shown that the optimal solution (29) resulted from differentiating the Lagrangian function (30) and satisfying the KKT conditions is a global solution. ■

Note that the structure of (29) is similar to the one in (23). However, (29) does not depend on γ_{vi}, as in (23). This is because we did not adopt the channel inversion power policy in the optimal approach. In addition to the effect of the optimized sensing threshold, γ_{ui}, the impact of the soft-sensing information parameter $\frac{h_0(\zeta_i)}{h_1(\zeta_i)}$ appears in (29). As stated earlier, in the explanation of (18), increasing $\frac{h_0(\zeta_i)}{h_1(\zeta_i)}$ leads to larger feasibility region of the interference constraint. Thus, enlarging the feasible set of the optimal P_{si}. Furthermore, $\frac{h_0(\zeta_i)}{h_1(\zeta_i)}$ directly decreases P_{si}, but, increases P_w, which also affect P_{si} indirectly.

The optimal value of the variable t is derived for general N as in (15). In particular, for $N = 1$, t is derived as follows,

$$t^* = \left[\frac{\exp(1/\lambda_1)}{\gamma_{si}\lambda_1} (1 + P_p\gamma_{psi}) \left(k_t\lambda_1 + \overline{\gamma_{spi}}\lambda_2 \left[\alpha K_{01}(\gamma_{ui}) + \overline{\alpha} K_{00}(\gamma_{ui}) \frac{h_0(\zeta_i)}{h_1(\zeta_i)} \right] \right) \right] - \frac{\lambda_3}{\lambda_1}. \quad (31)$$

Note that (31) has a similar structure as (28), except for the effect of the sensing threshold γ_{ui}.

IV. Simulation Results

In this section we begin by finding the expressions of the previously derived formulas based on specific distribution. Furthermore, we numerically evaluate both the sub-optimal and the optimal proposed resource allocation schemes in comparison with the benchmark ones.

A. Application on Specific Distributions

We assume that the fading channels γ_{si}, γ_{spi}, and γ_{psi} are independent identically distributed random variables that follow an exponential distribution. Furthermore, although ζ_i is not a Gaussian random variable, we use the central limit theorem to approximate it as a Gaussian distributed random variable, which simplifies the analysis without loss of generality. This approximation is done for both hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, with means μ_0, μ_1, and variances δ_0^2, δ_1^2, respectively, and can be calculated as in [26]. Since we obtain a sufficiently large number of samples for ζ_i, it is reasonable to use the Gaussian approximation in our analysis. The corresponding PDF of both hypotheses are $h_0(\zeta_i)$ and $h_1(\zeta_i)$, and expressed as, $h_0(\zeta_i) = \frac{\exp(-\frac{(\zeta_i-\mu_0)^2}{2\delta_0^2})}{\sqrt{2\pi}\delta_0}$, and $h_1(\zeta_i) = \frac{\exp(-\frac{(\zeta_i-\mu_1)^2}{2\delta_1^2})}{\sqrt{2\pi}\delta_1}$.

Following the same assumption on ζ_i's PDF, we can find that the region $Z_0 \in [\rho_1(\gamma_{ui}), \rho_2(\gamma_{ui})]$, which can be found as,

$$\rho_1(\gamma_{ui}) = a_0 \left(a_1 + \sqrt{b - \log(\gamma_{ui})}\right) \quad \text{and} \quad \rho_2(\gamma_{ui}) = a_0 \left(a_1 - \sqrt{b - \log(\gamma_{ui})}\right), \tag{32}$$

where, $a_0 = \frac{\sqrt{2\delta_1^2\delta_0^2(\delta_1^2-\delta_0^2)}}{\delta_1^2-\delta_0^2}$, $a_1 = \frac{\delta_1^2\mu_0-\delta_0^2\mu_1}{\sqrt{2\delta_1^2\delta_0^2(\delta_1^2-\delta_0^2)}}$, and $b = \frac{(\mu_1-\mu_0)^2}{2(\delta_1^2-\delta_0^2)} + \log\left(\sqrt{\frac{\delta_1^2}{\delta_0^2}}\right)$. Following similar assumptions, (14) can be simplified as,

$$\gamma_{vi} = W\left(\frac{\lambda_{2i}\overline{\gamma_{spi}}}{\lambda_{4i}}\sigma_i\right), \tag{33}$$

where W is the principal branch of the Lambert function [27], which is the solution of $x\exp(x) = y$, i.e., $x = W(y)$.

Furthermore, exact values can be found for the mentioned general functions of both γ_{vi} and γ_{ui}, as follows,

$$G(\gamma_{vi}) = E_1(\gamma_{vi}), \tag{34}$$

where $E_1(.)$ is the exponential integral function given by $E_1(x) = \int_x^\infty \frac{e^{-t}}{t}dt$. Also, we find that

$K_{01}(\gamma_{ui})$ can be expressed as follows,

$$K_{01}(\gamma_{ui}) = Q(f_{12}(\gamma_{ui})) - Q(f_{11}(\gamma_{ui})), \tag{35}$$

where Q(.) is the Q-function, defined as $Q(x) = \frac{1}{2\pi}\int_x^\infty e^{-\frac{t^2}{2}}dt$. The function $f_{11}(\gamma_{ui}) = \frac{\rho_1(\gamma_{ui})-\mu_1}{\delta_1}$, and $f_{12}(\gamma_{ui}) = \frac{\rho_2(\gamma_{ui})-\mu_1}{\delta_1}$. Furthermore, $K_{00}(\gamma_{ui}) = Q(f_{02}(\gamma_{ui}))-Q(f_{01}(\gamma_{ui}))$, where, $f_{01}(\gamma_{ui}) = \frac{\rho_1(\gamma_{ui})-\mu_0}{\delta_0}$, and $f_{02}(\gamma_{ui}) = \frac{\rho_2(\gamma_{ui})-\mu_0}{\delta_0}$.

In the next sub-section, numerical evaluation of all systems are presented under the assumptions mentioned in this sub-section.

B. Numerical Evaluation of both the Proposed and Benchmark Systems.

In this section we numerically evaluate the performance of the proposed systems (optimal and sub-optimal schemes) versus the corresponding benchmark systems (optimal and sub-optimal scheme without sensing information). Two metrics are considered for the system evaluation. One is the secondary minimum achievable EPG. The second is the outage metric, which is defined as the event where there is no feasible solution of the corresponding optimization problem. The simulation parameters are mentioned in Table I.

TABLE I
SIMULATION PARAMETERS

Parameter Name	**Value**
Sensing averaging bits	100
Primary transmitter power (P_p)	27 dBm
Secondary transmitter power (P_{max})	27 dBm [28]
# Monte-Carlo iterations	10000
Wireless channels	Rayleigh, Slow Flat Fading
# Subcarriers (N)	8
P (fixed case)	20 dBm, 27 dBm
Q_{int} (fixed case)	0 dBm, 5 dBm
Minimum Rate (R_{min})	1 bit / symbol

It is known that the minimum EPG is resulted from minimizing the secondary transmission power and the corresponding rate. However, note that we have a minimum rate constraint R_{min}.

Therefore, we define the EPG outage probability as follows,

$$\text{EPG Outage} = \begin{cases} 0 & ; \text{ A feasible solution exists} \\ 1 & ; \text{ No feasible solution exists} \end{cases}. \tag{36}$$

Figure 3 shows the EPG outage performance of all proposed and benchmark schemes versus P_{max}. It is observed that both proposed schemes outperform the benchmark ones. We also note that the difference in performance between the proposed optimal scheme and the proposed sub-optimal scheme depends on other parameters, such as, d_s and N, where, d_s is the distance between the ST and SR. Note that at low $d_s = 5$m both scheme achieve almost similar performance. Increasing d_s up to 12m increases the gap between both schemes such that the optimal scheme achieves 0.1 lower outage performance.

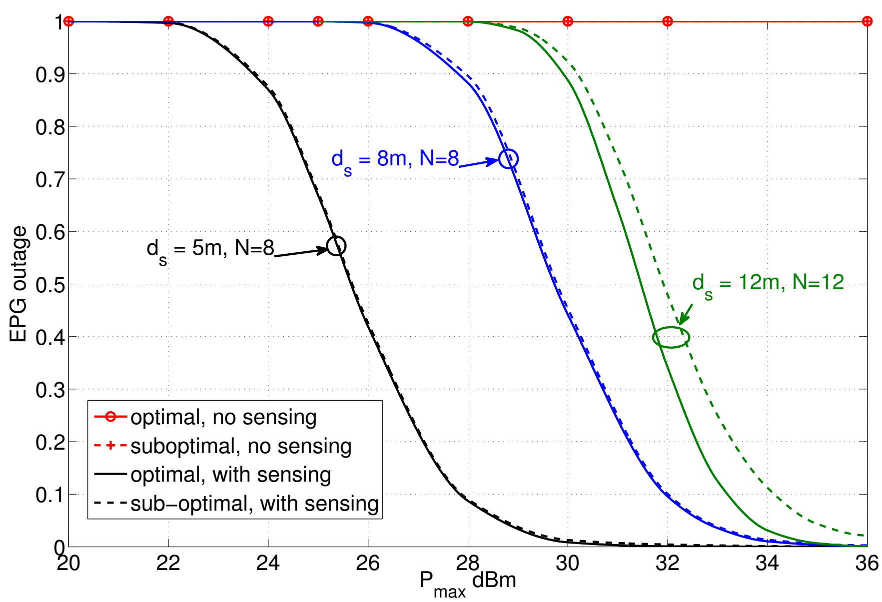

Fig. 3. EPG outage performance versus P_{max} for all schemes.

Figure 4 shows the CDF performance of the EPG metric for different P_{max}. We note the slight improvement of the optimal scheme compared to the sub-optimal scheme. It is observed that by increasing P_{max}, up to a certain value, the CDF performance of the EPG metric improves for both proposed schemes.

Figure 5 shows the EPG outage performance of both the proposed schemes, optimal and

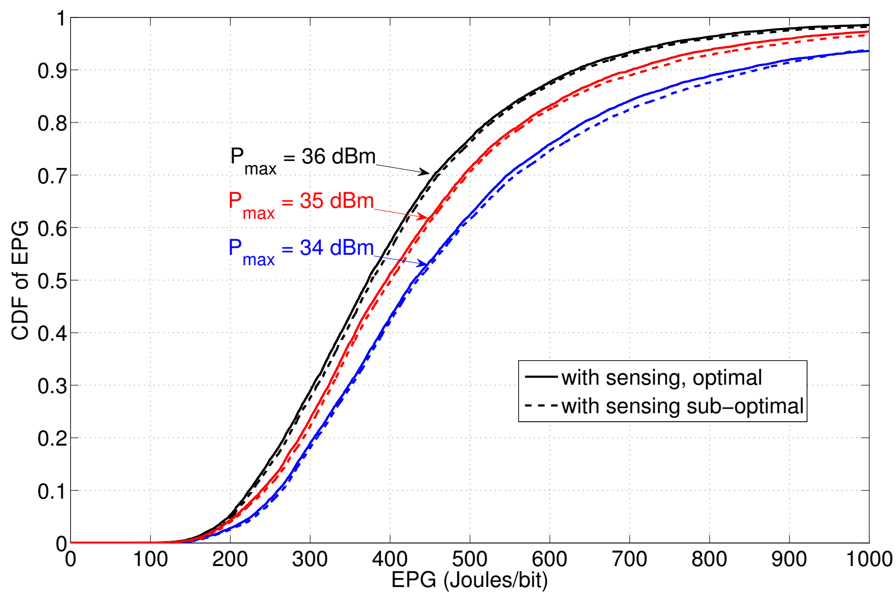

Fig. 4. CDF performance of the EPG metric for different P_{max} for all schemes, $d_s = 12$ m.

sub-optimal, versus the number of subcarriers (N) for two values of maximum power threshold (P_{max}). It is observed that the optimal scheme outperforms the sub-optimal one under both $P_{max} = 27$ dBm and $P_{max} = 29$ dBm. However, we note that by decreasing the power threshold (P_{max}) the difference in performance increases between both schemes. Such that by decreasing P_{max} from 29 dBm to 27 dBm, the difference in outage performance, between both schemes, increases from 0.02 up to 0.22. This means that the performance improvement of the optimal schemes is better highlighted at low power thresholds. As expected, the performance of both schemes improves by increasing the number of sub-carriers N.

Figure 6 shows the CDF performance of the EPG metric for both proposed schemes, optimal and sub-optimal, under $N = 36$ and $N = 46$. It is observed that the optimal CDF performance of the EPG metric outperforms the performance of the sub-optimal scheme in all cases. It is noted that the gap in the performance improvement increases by increasing N from 36 to 46.

Figure 7 shows the EPG outage performance of the optimal scheme, proposed and benchmark, versus α for $N = 4$ and $N = 6$. It is observed that the outage performance of the proposed scheme degrades by increasing α, as expected. On the other hand, the outage performance of the benchmark scheme does not change by changing α. Note that the outage performance does not reach zero at $\alpha \to 0$ because of the sensing metric effect in the optimal scheme. This occurs

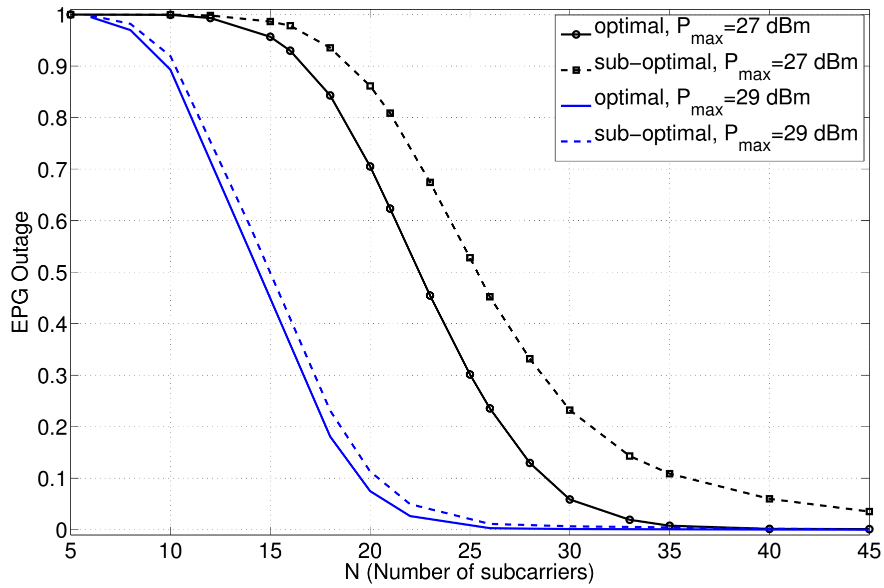

Fig. 5. EPG outage performance versus N for both proposed schemes, $d_s = 7$ m.

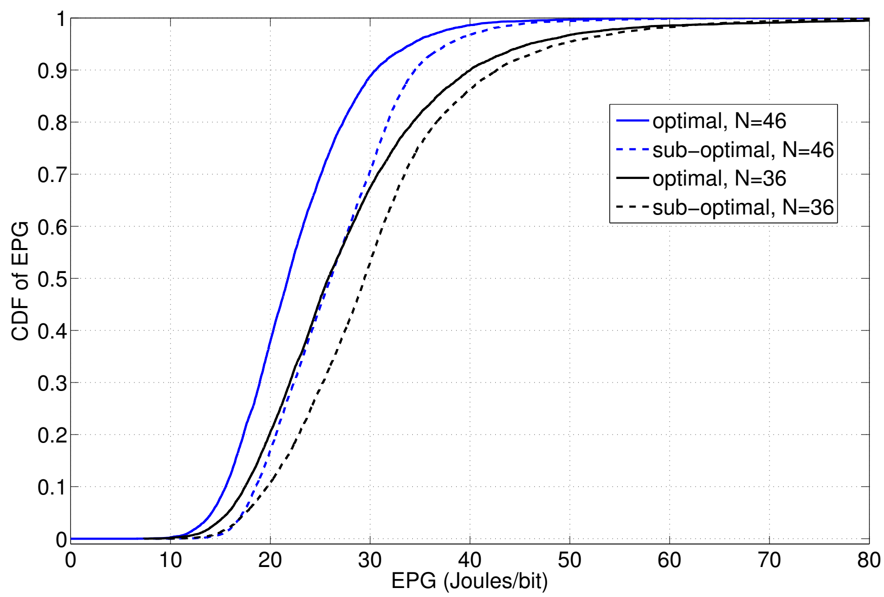

Fig. 6. CDF performance of the EPG metric for different N for all proposed schemes, $d_s = 7$ m.

because we always have error in the sensing performance.

Figure 8 shows the CDF performance of the EPG metric for different values of α under the optimal proposed scheme. It is observed that the CDF performance improves by decreasing the value of α. We note that at low values of α the CDF performance does not achieve a large improvement gap as in the case of large α. For example, at high values of α the improvement

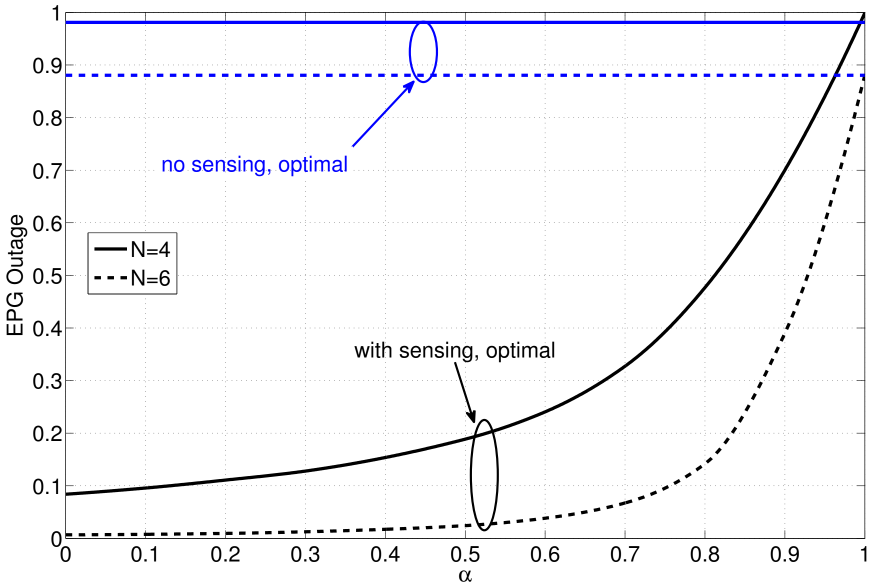

Fig. 7. EPG outage performance versus α for the optimal proposed and benchmark schemes, $d_s = 1$ m.

is noticeable, about 15 EPG units by decreasing α from $\alpha = 1$ to $\alpha = 0.9$. However, decreasing α from $\alpha = 0.1$ to $\alpha = 0$ does not achieve high improvement.

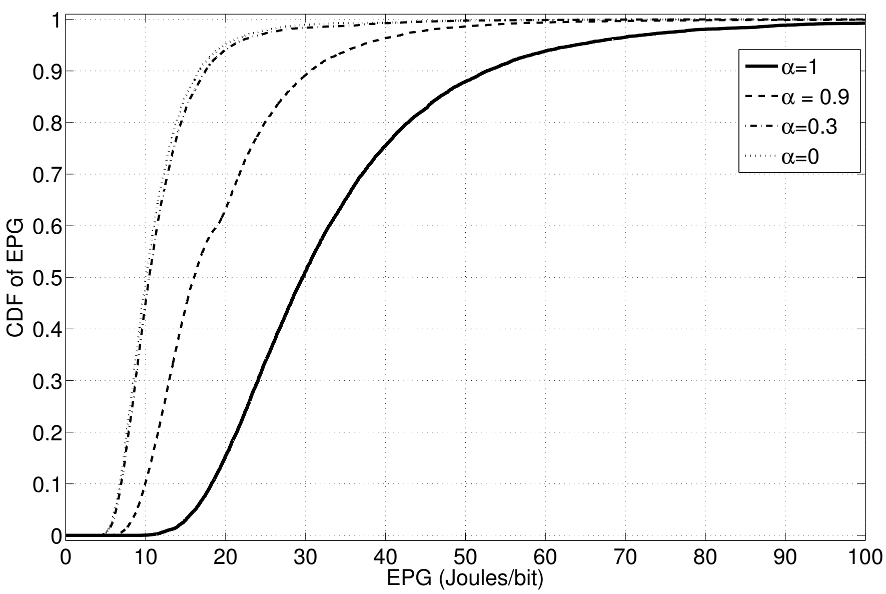

Fig. 8. CDF performance of the EPG metric for different α for the optimal proposed scheme, $d_s = 1$ m.

Figure 9 shows the EPG outage performance of the optimal schemes, proposed and benchmark, versus the number of sensing samples (N_s). It is seen that by increasing N_s the outage performance of the proposed scheme improves until it saturates on a certain value. It is observed that changing N_s does not affect the benchmark schemes. Note that the improvement gap increases by increasing N, which is observed for $N = 4$ and $N = 8$.

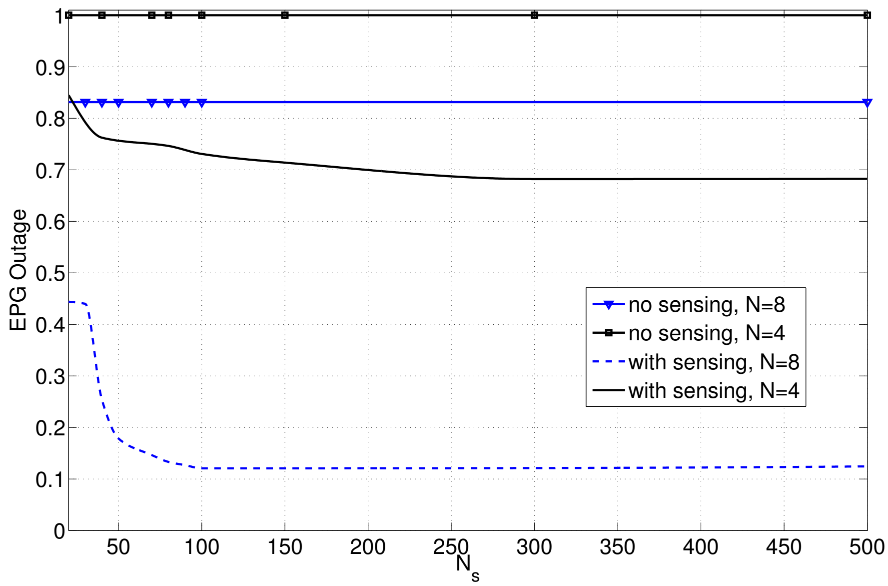

Fig. 9. EPG outage performance versus N_s for the optimal proposed and benchmark schemes, $d_s = 2$ m.

Three new systems, other than the benchmark and proposed systems, have been developed to extensively evaluate the proposed system. The evaluation of these systems are shown in Fig. 10 and their assumptions are described as follows. The first system minimizes the EPG metric with a fixed sensing threshold. This system does not optimize over the sensing threshold, γ_{ui}, it is noted in Fig. 10 as “Min EPG, FST”. The second system minimizes the power metric under similar constraints. It is noted in Fig. 10 as “Min Ps”. The third system considers the cognitive radio network with the PU being ideal all the time. This system is noted in Fig. 10 as “Min EPG, PI”. Finally, the original proposed system, optimal approach, with soft-sensing information is noted Fig. 10 as “Min EPG, SS”. Note that all systems utilize the onboard sensor. Figure 10 shows the EPG CDF of the normalized EPG. The normalization is done with respect to the maximum EPG value. It is observed that the “PI” achieves a slightly better results compared to the “SS” system. Note that for the same probability (around 0.95) the proposed system, “SS”, outperforms the other two schemes “FST” and “Min Ps” by about 0.05 and 0.25 normalized

EPG units, respectively.

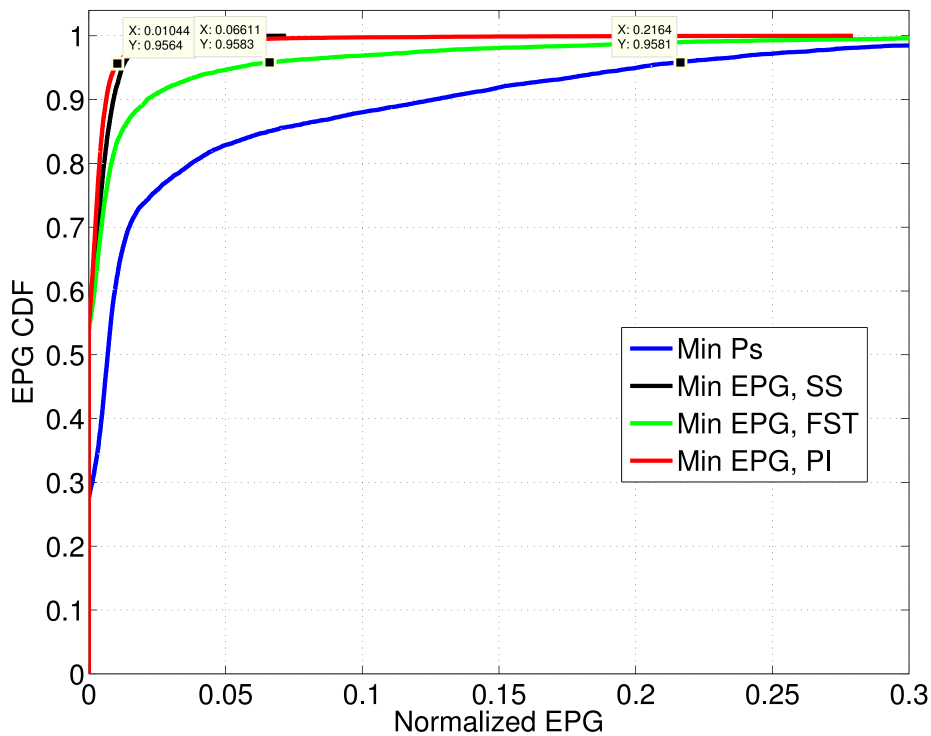

Fig. 10. Numerical Evaluation of all systems, Minimum EPG with soft-sensing (proposed system) "SS", fixed sensing threshold "FST", idle primary user "PI", and the minimum power problem "Min Ps", α = 0.1, $P_s = P_p = 25$ dBm, R_{min} = 0.5 bits per channel use.

V. Conclusion

In this paper we proposed two energy efficient resource allocation schemes that utilize the sensing information. The sub-optimal scheme is proposed based on the channel inversion power policy, whereas the optimal scheme is based on the calculus of variation principle. We analyzed both schemes, and derived their performances. Also, we analyzed the benchmark systems where no sensing is used, for both schemes to evaluate the improvement of the proposed ones. In addition to our analysis, we showed the superiority of the proposed systems over the benchmark systems through numerical results. This improvement reaches up to 20 Joules/bit reduction under the proposed system case compared to the benchmark one. The proposed optimal scheme reaches a reduction of up to 5 Joules/bit in comparison with the sub-optimal proposed scheme. Utilizing numerical simulations, we compared the optimal scheme with three other benchmarks and showed the relative improvements.

Appendix A
Proof of Lemma 1

To proof this lemma, it is necessary to verify that the objective function of problem $\mathcal{P}_0$ is pseudo-convex function with respect to the optimization variables, $\mathbf{P_s}(\boldsymbol{\gamma})$. The corresponding constraints must be verified to be quasi-convex ones with respect to the optimization variables, $\mathbf{P_s}(\boldsymbol{\gamma})$. It is clear that all the constraint are convex ones. Consequently, they are quasi-convex constraints. To show that $\mathcal{E}(\mathbf{P_s}(\boldsymbol{\gamma}))$, (4a), is pseudo-convex, two conditions must be satisfied [24],

1) $\mathcal{E}(\mathbf{P_s}(\boldsymbol{\gamma}))$ is a quasi-convex function on its domain $\mathcal{S}$.
2) If $\mathbf{P}^*_\mathbf{s}(\boldsymbol{\gamma}) \in \mathcal{S},\ \nabla\mathcal{E}(\mathbf{P}^*_\mathbf{s}(\boldsymbol{\gamma})) = 0$, then $\mathbf{P}^*_\mathbf{s}(\boldsymbol{\gamma})$ is a local minimum for $\mathcal{E}$.

The definition of a quasi-convex function is stated as follows, $f(\lambda x_1+(1-\lambda)x_2) \leq \max(f(x_1), f(x_2))$. To show the quasi-convexity structure of $\mathcal{E}(\mathbf{P_s}(\boldsymbol{\gamma}))$, we divide it into numerator and denominator terms as, $\mathcal{E}(\mathbf{P_s}(\boldsymbol{\gamma})) = \frac{\mathcal{E}_n(\mathbf{P_s}(\boldsymbol{\gamma}))}{\mathcal{E}_d(\mathbf{P_s}(\boldsymbol{\gamma}))} = \frac{k_t\sum_{i=1}^{N} P_{si}(\boldsymbol{\gamma})+k_c}{\sum_{i=1}^{N}\log(1+\frac{P_{si}(\boldsymbol{\gamma})\gamma_{si}}{1+P_p\gamma_{psi}})}$. Convexity and concavity properties of each $\mathcal{E}_n$ and $\mathcal{E}_d$ are utilized to show the quasi-convexity of $\mathcal{E}$, as follows (for notation simplicity we alternately use equivalent notations as $\mathbf{P_s}(\boldsymbol{\gamma}) = \mathbf{P_s}$, $\mathcal{E}(\boldsymbol{\gamma}) = \mathcal{E}$, $\mathcal{E}_\mathbf{n}(\boldsymbol{\gamma}) = \mathcal{E}_\mathbf{n}$, and $\mathcal{E}_\mathbf{d}(\boldsymbol{\gamma}) = \mathcal{E}_\mathbf{d}$, we assume that $\mathcal{E}(\mathbf{P}^{(\mathbf{2})}_\mathbf{s}) \leq \mathcal{E}(\mathbf{P}^{(\mathbf{1})}_\mathbf{s})$),

$$\begin{aligned}
\mathcal{E}_n(\lambda\mathbf{P}^{(\mathbf{1})}_\mathbf{s} + (1-\lambda)\mathbf{P}^{(\mathbf{2})}_\mathbf{s}) &= \lambda\mathcal{E}_n(\mathbf{P}^{(\mathbf{1})}_\mathbf{s}) + (1-\lambda)\mathcal{E}_n(\mathbf{P}^{(\mathbf{2})}_\mathbf{s}) && \text{(37a)}\\
&\leq \lambda\mathcal{E}_n(\mathbf{P}^{(\mathbf{1})}_\mathbf{s}) + (1-\lambda)\frac{\mathcal{E}_n(\mathbf{P}^{(\mathbf{1})}_\mathbf{s})}{\mathcal{E}_d(\mathbf{P}^{(\mathbf{1})}_\mathbf{s})}\mathcal{E}_d(\mathbf{P}^{(\mathbf{2})}_\mathbf{s}) && \text{(37b)}\\
&= \frac{\mathcal{E}_n(\mathbf{P}^{(\mathbf{1})}_\mathbf{s})}{\mathcal{E}_d(\mathbf{P}^{(\mathbf{1})}_\mathbf{s})}\left[\lambda\mathcal{E}_d(\mathbf{P}^{(\mathbf{1})}_\mathbf{s}) + (1-\lambda)\mathcal{E}_d(\mathbf{P}^{(\mathbf{2})}_\mathbf{s})\right] && \text{(37c)}\\
&\leq \frac{\mathcal{E}_n(\mathbf{P}^{(\mathbf{1})}_\mathbf{s})}{\mathcal{E}_d(\mathbf{P}^{(\mathbf{1})}_\mathbf{s})}\left[\mathcal{E}_d\left(\lambda\mathbf{P}^{(\mathbf{1})}_\mathbf{s} + (1-\lambda)\mathbf{P}^{(\mathbf{2})}_\mathbf{s}\right)\right] && \text{(37d)}\\
&\implies \frac{\mathcal{E}_n\left(\lambda\mathbf{P}^{(\mathbf{1})}_\mathbf{s} + (1-\lambda)\mathbf{P}^{(\mathbf{2})}_\mathbf{s}\right)}{\mathcal{E}_d\left(\lambda\mathbf{P}^{(\mathbf{1})}_\mathbf{s} + (1-\lambda)\mathbf{P}^{(\mathbf{2})}_\mathbf{s}\right)} \leq \frac{\mathcal{E}_n(\mathbf{P}^{(\mathbf{1})}_\mathbf{s})}{\mathcal{E}_d(\mathbf{P}^{(\mathbf{1})}_\mathbf{s})} && \text{(37e)}\\
&\implies \mathcal{E}(\lambda\mathbf{P}^{(\mathbf{1})}_\mathbf{s} + (1-\lambda)\mathbf{P}^{(\mathbf{2})}_\mathbf{s}) \leq \mathcal{E}(\mathbf{P}^{(\mathbf{1})}_\mathbf{s}), && \text{(37f)}
\end{aligned}$$

where (37b) results from the fact that $\mathcal{E}(\mathbf{P_s^{(2)}}) \leq \mathcal{E}(\mathbf{P_s^{(1)}})$, (37d) is valid because of the concave property of $\mathcal{E}_d$.

The second step, in of proving the pseudo-convexity of $\mathcal{E}\left(\mathbf{P_s}\right)$, is to show that $\nabla\mathcal{E}\left(\mathbf{P_s^*}\right) = 0$. It is clear that $\nabla\mathcal{E}\left(\mathbf{P_s^*}\right) = 0$ iff $\nabla\mathcal{E}_n\left(\mathbf{P_s^*}\right)\mathcal{E}_d\left(\mathbf{P_s^*}\right) - \mathcal{E}_n\left(\mathbf{P_s^*}\right)\nabla\mathcal{E}_d\left(\mathbf{P_s^*}\right) = 0$. Then, $\nabla\mathcal{E}_n\left(\mathbf{P_s^*}\right) = \mathcal{E}\left(\mathbf{P_s^*}\right)\nabla\mathcal{E}_d\left(\mathbf{P_s^*}\right)$. Utilizing the properties of $\mathcal{E}_n$ function,

$$\begin{aligned}
\mathcal{E}_n\left(\mathbf{P_s}\right) &= \mathcal{E}_n\left(\mathbf{P_s^*}\right) + \nabla\mathcal{E}_n\left(\mathbf{P_s^*}\right)\left(\mathbf{P_s} - \mathbf{P_s^*}\right) && \text{(38a)}\\
&= \mathcal{E}_n\left(\mathbf{P_s^*}\right) + \mathcal{E}\left(\mathbf{P_s^*}\right)\nabla\mathcal{E}_d\left(\mathbf{P_s^*}\right)\left(\mathbf{P_s} - \mathbf{P_s^*}\right) && \text{(38b)}\\
&\geq \mathcal{E}_n\left(\mathbf{P_s^*}\right) + \mathcal{E}\left(\mathbf{P_s^*}\right)\left(\mathcal{E}_d\left(\mathbf{P_s}\right) - \mathcal{E}_d\left(\mathbf{P_s^*}\right)\right) && \text{(38c)}\\
&= \mathcal{E}_n\left(\mathbf{P_s^*}\right) + \mathcal{E}\left(\mathbf{P_s^*}\right)\mathcal{E}_d\left(\mathbf{P_s}\right) - \mathcal{E}_n\left(\mathbf{P_s^*}\right) && \text{(38d)}\\
&\implies \frac{\mathcal{E}_n\left(\mathbf{P_s}\right)}{\mathcal{E}_d\left(\mathbf{P_s}\right)} \geq \mathcal{E}\left(\mathbf{P_s^*}\right), && \text{(38e)}
\end{aligned}$$

where (38a) follows from the convex property of $\mathcal{E}_n$ and (38c) is due to the concave structure of $\mathcal{E}_d\left(\mathbf{P_s}\right)$.

By combining (37) and (38), it is clear that the function $\mathcal{E}\left(\mathbf{P_s}\right)$ is pseudo-convex. Utilizing Theorem 1, the global optimal solution of problem $\mathcal{P}_0$ is obtained by satisfying the KKT conditions.

APPENDIX B
PROOF OF LEMMA 2

Without loss of generality, we select $\mathbf{x}_1 \leq \mathbf{x}_2$. By definition, $g(\mathbf{x}_2) \leq g(\mathbf{x}_1)$. Under the assumption that $z(\mathbf{x}_2) \leq z(\mathbf{x}_1)$ we have two cases for $f(\mathbf{x})$. Case 1, where $f(\mathbf{x}_2) \leq f(\mathbf{x}_1)$, thus, $f(\lambda x_1 + (1-\lambda)x_2) \leq f(\mathbf{x}_1)$. Case 2, where $f(\mathbf{x}_1) \leq f(\mathbf{x}_2)$, thus, $f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leq f(\mathbf{x}_2)$.

1) Case 1, $f(\mathbf{x}_2) \leq f(\mathbf{x}_1)$: To prove the quasi-convexity we need to show that $z(\lambda x_1 + (1-\lambda)x_2) < z(\mathbf{x}_1)$ as follows,

$$\begin{aligned}
f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) &\leq f(\mathbf{x}_1)\frac{g(\mathbf{x}_1)}{g(\mathbf{x}_1)} && \text{(39a)}\\
&\leq f(\mathbf{x}_1)\frac{g(\mathbf{x}_1)}{g(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2)} && \text{(39b)}
\end{aligned}$$

$$\implies z(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leq z(\mathbf{x}_1), \tag{39c}$$

where (39a) follows from the quasi-convex definition, and (39b) results from the fact that $g(\mathbf{x})$ is a non-increasing function.

2) Case 2, $f(\mathbf{x}_1) \leq f(\mathbf{x}_2)$: Knowing that $f(\mathbf{x}_1) \leq f(\mathbf{x}_2)$ and $g(\mathbf{x}_2) \leq g(\mathbf{x}_1)$ it is not always clear that $z(\mathbf{x}_2) \leq z(\mathbf{x}_1)$. Therefore, we adopt different approach as follows,

$$g(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leq g(\mathbf{x}_1) \tag{40a}$$

$$\leq g(\mathbf{x}_1)\frac{f(\mathbf{x}_1)}{f(\mathbf{x}_2)} \tag{40b}$$

$$\leq g(\mathbf{x}_1)\frac{f(\mathbf{x}_1)}{f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2)} \tag{40c}$$

$$\implies z(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leq z(\mathbf{x}_1), \tag{40d}$$

where (40c) follows from the definition of $f(\mathbf{x})$.

Combining both case 1 and case 2, $\forall\ z(\mathbf{x}_2) \leq z(\mathbf{x}_1) \implies z(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leq z(\mathbf{x}_1)$. This verifies that $z(\mathbf{x})$ is a quasi-convex function.

APPENDIX C
THE EQUIVALENCE BETWEEN TWO PROBLEMS.

The equivalence between problem $\mathcal{P}_1'$ and problem $\mathcal{P}_1$ is proved in two steps, one is showing that the optimal solution of $\mathcal{P}_1$ is optimal for $\mathcal{P}_1'$, whereas, the other is showing that the optimal solution of $\mathcal{P}_1'$ is optimal for $\mathcal{P}_1$. First, let $\mathbf{P}_\mathbf{s}^*$ be an optimal solution of $\mathcal{P}_1$ that minimizes $\mathcal{E} = \frac{k_t \sum_{i=1}^{N} P_{si}(\gamma)+k_c}{\sum_{i=1}^{N}\log(1+\frac{P_{si}(\gamma)\gamma_{si}}{1+P_p\gamma_{psi}})}$ under (4b), (4c), and (8). It follows that $\mathbf{P}_\mathbf{s}^*$ is a feasible solution for $\mathcal{P}_1$ and $\mathcal{P}_1'$ since it does not violate any of their constraints. Let us assume that it is possible to have an optimal solution for $\mathcal{P}_1'$ that is better than $\mathbf{P}_\mathbf{s}^*$, then the EPG metric in constraint (13b), of problem $\mathcal{P}_1'$, is larger than the EPG metric (also objective function) of problem $\mathcal{P}_1$. This means that this solution either violates one of the constraints (13c), (13d), (13e), and (13f) or $\mathbf{P}_\mathbf{s}^*$ is not an optimal solution of problem $\mathcal{P}_1$ which contradicts with the above assumption (that $\mathbf{P}_\mathbf{s}^*$ is the optimal solution of $\mathcal{P}_1$). Second, assume that the optimal pair of the variables (t , $\mathbf{P}_\mathbf{s}$) that optimizes problem $\mathcal{P}_1'$ is designated as (t^* , $\mathbf{p}_1'^*$) and the corresponding optimal EPG value is designated as EPG*. Note that minimizing t, the objective function of $\mathcal{P}_1'$, decreases the left hand side of constraint (13b), which is the EPG, up to a certain value *such that*, EPG$^* = t^*$. Note that the optimal $\mathbf{p}_1'^*$ satisfies all the other constraints (13c), (13d), (13e), and (13f), and they do not depend on t. Since $\mathbf{p}_1'^*$ minimizes EPG under all constraints of $\mathcal{P}_1'$, thus all the constraints of $\mathcal{P}_1$, then $\mathbf{p}_1'^*$ minimizes the EPG metric in $\mathcal{P}_1$ under the corresponding constraints. Now, let us assume there exist an optimal solution for $\mathcal{P}_1$ that is better than $\mathbf{p}_1'^*$. This solution either violates $\mathcal{P}_1$'s constraints (4b), (4c), and (8) or the optimal pair (t^*,$\mathbf{p}_1'^*$) does not minimizes the EPG of $\mathcal{P}_1'$ and the corresponding objective function t, which contradicts with the above assumption. Finally, we conclude from the first and the second steps that problem $\mathcal{P}_1'$ is equivalent to $\mathcal{P}_1$.

Appendix D
Proof of Theorem 2

We will separate the proof into three parts corresponding to each of the optimization variables σ_i, γ_{vi}, and t respectively.

The first part is to prove part (1) of Theorem 2, we generate the Lagrangian function of problem $\mathcal{P}'_1$ [25], [29], by including both (13)'s objective function and inequalities constraint, as part of the Lagrangian function. We leave the maximum power constraint (13c) to the end as an upper-bound, which can also be called clipping policy, just like in the case of iterative water-filling algorithm, as follows,

$$L = t + \lambda_1 \left[k_t \sum_{i=1}^{N} \frac{\sigma_i}{\gamma_{si}} + k_c - t \sum_{i=1}^{N} \log(1 + \frac{\sigma_i}{1 + P_p \gamma_{psi}}) \right] - \lambda_3 \left[\sum_{i=1}^{N} \log(1 + \frac{\sigma_i}{1 + P_p \gamma_{psi}}) - R_{min} \right] + \sum_{i=1}^{N} \lambda_{2i} \left[\overline{\gamma_{spi}} \sigma_i G(\gamma_{vi}) - Q_{int} \right] + \sum_{i=1}^{N} \lambda_{4i} \left[\gamma_{vi} - \gamma_{si} \right]. \tag{41}$$

Differentiating the Lagrangian function w.r.t. σ_i and equating to zero, $\frac{\partial L}{\partial \sigma_i} = 0$, we obtain the following general solution for σ_i,

$$\frac{\sigma_i^*}{\gamma_{si}} = \left[\frac{\lambda_1 t + \lambda_3}{k_t \lambda_1 + \lambda_{2i} \left[\overline{\gamma_{spi}} \sigma_i G(\gamma_{vi}) - Q_{int} \right]} - \frac{1 + P_p \gamma_{psi}}{\gamma_{si}} \right]^+, \tag{42}$$

where, utilizing the Complementary Slackness conditions for constraint (12) we have two cases to solve for σ_i:

- $\lambda_{2i} = 0$, then we obtain $\frac{\sigma_i}{\gamma_{si}} = \min\left(\left[\frac{\lambda_1 t+\lambda_3}{k_t\lambda_1} - \frac{1+P_p\gamma_{psi}}{\gamma_{si}}\right]^+, P_{max}\right)$.
- $\lambda_{2i} \neq 0$, from the slackness conditions we know that (12) becomes equality, then $G(\gamma_{vi}) = \frac{Q_{int}}{\sigma_i\overline{\gamma_{spi}}})$, and σ_i is obtained by solving the following quadratic equation

$$\frac{\sigma_{i,*}^2}{\gamma_{si}} + \sigma_{i,*}\left[\frac{1+\gamma_{psi}P_p}{\gamma_{si}} + \frac{\lambda_{2i}}{\lambda_1}Q_{int} - \frac{(\lambda_1 t+\lambda_3)}{k_t\lambda_1}\right] + \left[\frac{\lambda_{2i}}{k_t\lambda_1}Q_{int}(1+\gamma_{psi}P_p)\right] = 0, \quad (43)$$

where $\sigma_i^* = \min\left(\sigma_{i,*}, \sigma_{mn\sigma_{mn}s}\right)$

where, σ_{mn} is defined as in Theorem 2.

Then, in order to prove the second part of Theorem 2, it is direct to see that the optimal γ_{vi} can be found by taking partial derivative of (41) w.r.t. γ_{vi} and equating it to zero, as follows,

$$\frac{\partial}{\partial\gamma_{vi}}L = \frac{\partial}{\partial\gamma_{vi}}G(\gamma_{vi})\lambda_{2i}\overline{\gamma_{spi}}\sigma_i + \lambda_{4i} = 0. \quad (44)$$

from which (14) is immediate. Finally, the third part of Theorem 2, optimal value of t, can be derived using the same Lagrangian function in (41) and finding the zeros of $\frac{\partial L}{\partial t} = 0$.

Appendix E

Structure of The Interference Constraint.

We begin by proving the monotonicity of constraint (18) on the variable γ_{ui}. Then, the non-increasing property of (18) is verified with respect to variable γ_{ui}.

It is enough to prove that the function $H(\gamma_{ui})$, defined below, is a monotonic function, on γ_{ui}, to guarantee that constraint (18) is a monotonic one, on γ_{ui}.

$$H(\gamma_{ui}) = \frac{\alpha K_{01}(\gamma_{ui}) + \overline{\alpha}K_{00}(\gamma_{ui})}{\alpha Q_{int}K_{01}(\gamma_{ui}) + \overline{\alpha}PK_{00}(\gamma_{ui})}. \quad (45)$$

We ignored both σ_i and $G(\gamma_{vi})$ from the constraint because both of them does not depend on γ_{ui}, so they are just a scaling factor. Note that both $K_{01}(\gamma_{ui}) = \int_{Z_0} h_1(\zeta_i)d\zeta_i$ and $K_{00}(\gamma_{ui}) = \int_{Z_0} h_0(\zeta_i)d\zeta_i$ are positive quantities. It is assumed that $h_1(\zeta_i)$ has mean μ_1 and variance δ_1^2 which are distant from the mean μ_0 and variance δ_0^2 of $h_0(\zeta_i)$. Note that $H(\gamma_{ui})$ is monotonic function if it does not change from increasing to decreasing (or vice versa). Since $H(\gamma_{ui})$ is assumed to be continuous differentiable over the domain of γ_{ui}, then the monotonicity can be proved by proving that $\frac{\partial H(\gamma_{ui})}{\partial \gamma_{ui}} \neq 0$. This means that $H(\gamma_{ui})$ does not have a stationary point (local minimum or local maximum). We assume that $P >> Q_{int}$ [30], [31]. Then, it is clear that $H(\gamma_{ui}) \approx \frac{\alpha K_{01}(\gamma_{ui})}{\overline{\alpha}\frac{P}{Q_{int}}K_{00}(\gamma_{ui})} + \frac{1}{P}$. Note that $H(\gamma_{ui})$ is monotonic on γ_{ui} if $\frac{K_{01}(\gamma_{ui})}{K_{00}(\gamma_{ui})}$ is monotonic on γ_{ui}. Since both $K_{01}(\gamma_{ui})$ and $K_{01}(\gamma_{ui})$ are positive quantities, then it follows that if $H_n(\gamma_{ui}) = \frac{K_{01}(\gamma_{ui})}{K_{00}(\gamma_{ui})}$ is monotonic then $H(\gamma_{ui})$ is monotonic. Let us take the derivative of $H_n(\gamma_{ui})$ as follows,

$$\frac{\partial H_n(\gamma_{ui})}{\partial \gamma_{ui}} = \frac{K_{00}(\gamma_{ui})\frac{\partial K_{01}(\gamma_{ui})}{\partial \gamma_{ui}} - K_{01}(\gamma_{ui})\frac{\partial K_{00}(\gamma_{ui})}{\partial \gamma_{ui}}}{K_{00}(\gamma_{ui})}. \tag{46}$$

Then, $\frac{\partial H_n(\gamma_{ui})}{\partial \gamma_{ui}} = 0$ if,

$$K_{00}(\gamma_{ui})\frac{\partial K_{01}(\gamma_{ui})}{\partial \gamma_{ui}} - K_{01}(\gamma_{ui})\frac{\partial K_{00}(\gamma_{ui})}{\partial \gamma_{ui}} = 0. \tag{47}$$

We begin by assuming that (47) is satisfied and then prove that this assumption contradict with our original assumptions. Taking one term of (47) to the right hand side and taking integration for both sides results by the following,

$$\int \frac{1}{K_{01}(\gamma_{ui})}\partial K_{01}(\gamma_{ui}) = \int \frac{1}{K_{00}(\gamma_{ui})}\partial K_{00}(\gamma_{ui}) \implies \log\left(K_{01}(\gamma_{ui})\right) = \log\left(K_{00}(\gamma_{ui})\right). \tag{48}$$

Since $K_{01}(\gamma_{ui})$ and $K_{00}(\gamma_{ui})$ has distinct variance and mean parameters, then (48) cannot be true. By contradiction we prove that $\frac{\partial H(\gamma_{ui})}{\partial \gamma_{ui}} \neq 0$, which end the monotonicity proof of constraint (18).

In order to verify the non-increasing property of $H_n(\gamma_{ui})$ we consider that both $h_0(\zeta)$ and $h_1(\zeta)$ follows Gaussian distribution with mean and variance as mentioned in IV-A. Note that $\gamma_{ui} \geq \delta_1^2 Q^{-1}(P_D) + \mu_1$ is a lower bound on the sensing threshold γ_{ui} to guarantee a certain quality of detection probability. Therefore, it is clear that γ_{ui} is lower bounded by 1, since $\mu_1 \geq 1$. Recall that Z_0 is defined such that $\frac{h_0(\zeta)}{h_1(\zeta)} \geq \gamma_{ui}$. Thus, increasing γ_{ui} results in decreasing the part of $Z_0(\zeta)$ under $h_1(\zeta)$ and increasing the part of $Z_0(\zeta)$ under $h_1(\zeta)$. It follows that,

$$\forall \quad \gamma_{ui}^{(1)} \leq \gamma_{ui}^{(2)} \implies \begin{aligned} &\int_{Z_0^{(2)}} h_1(\zeta_i)d\zeta_i \leq \int_{Z_0^{(1)}} h_1(\zeta_i)d\zeta_i \\ &\int_{Z_0^{(2)}} h_0(\zeta_i)d\zeta_i \geq \int_{Z_0^{(1)}} h_0(\zeta_i)d\zeta_i, \end{aligned} \tag{49}$$

where, $Z_0^{(2)}$ corresponds to $\gamma_{ui}^{(2)}$ and $Z_0^{(1)}$ corresponds to $\gamma_{ui}^{(1)}$. It is clear from (49) that $H_n(\gamma_{ui})$

is a non-increasing function with respect to γ_{ui}.

APPENDIX F
BACKGROUND ON THE CALCULUS OF VARIATION

In this section, we provide the necessary background on the calculus of variation methodology. The reason for using channel inversion technique as a power policy is to avoid facing the difficult triple integrals in the sensing (average interference) constraint (16). However, we found that calculus of variation with some adjustments can solve the problem. It is known that calculus of variation is a powerful tool to solve optimization problems, which has integrals in their objective function. There are many books that deal with the theory of calculus of variations, as [32], [33], thus, we are not going to mention the proofs of the next mentioned theorems.

Suppose we want to find the function $y(x)$ that maximize or minimize the integral $\int_a^b F(y, y^\circ, x)$ w.r.t. the equality constraint $\varphi_k(y, y^\circ, x) = 0, \quad \forall k = 1, ..., K$, where, $y^\circ = \frac{\partial y}{\partial x}$, K is the number of constraints. Then we define a combination of the objective and constraint functions, a Lagrangian like function, as,

$$\widehat{F}(y, y^\circ, x) = F(y, y^\circ, x) + \sum_{k=1}^{K} \lambda_k(x)\varphi_k(y, y^\circ, x), \tag{50}$$

Finally, to get y, we solve the Euler-Lagrange equation,

$$\frac{\partial \widehat{F}}{\partial y} - \frac{d}{dx}\frac{\partial \widehat{F}}{\partial y^\circ} = 0, \tag{51}$$

where, the second term is zero in case $\widehat{F}$ does not depend on y°. To apply Euler-Lagrange equation on the inequality constraint instead of the equality, we add a new variable v^2 to the inequality constraint as follows,

$$\varphi_k(y, y^\circ, x) \leq 0 \implies \varphi_k(y, y^\circ, x) + v^2 = 0. \tag{52}$$

Now, we generalize Euler-Lagrange theorem into multiple dependent variables, functions, and multiple independent variables. We state the results of [32], section 17.3. Consider we have the following integral problem to be optimized,

$$J = \int_{x_{11}}^{x_{12}} \cdots \int_{x_{n1}}^{x_{n2}} f\left[y_1(x_1...x_n), ..., y_p(x_1...x_n), y_1^{\circ}(x_1...x_n), ..., y_p^{\circ}(x_1...x_n), x_1, ..., x_n\right] dx_1, ..., dx_n. \tag{53}$$

Considering the above integral problem, we adjust Euler-Lagrange as follows,

$$\frac{\partial f}{\partial y_i} - \sum_j \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial y_{ij}}\right) = 0, \forall \quad i = 1, 2, ..., p \quad \text{and} \quad \forall \quad j = 1, 2, ..., n, \tag{54}$$

where, $y_{ij} \equiv \frac{\partial y_i}{\partial x_j}$.

REFERENCES

[1] A. Anand, C. Manikopoulos, Q. Jones, and C. Borcea, "A quantitative analysis of power consumption for location-aware applications on smart phones," in *IEEE International Symposium on Industrial Electronics. ISIE*, 2007, pp. 1986–1991.

[2] C. Han, T. Harrold, and et al., "Green radio: radio techniques to enable energy-efficient wireless networks," *Communications Magazine, IEEE*, vol. 49, no. 6, pp. 46 –54, june 2011.

[3] Z. Hasan, H. Boostanimehr, and V. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Communications Surveys Tutorials*, vol. 13, no. 4, pp. 524 –540, quarter 2011.

[4] D. Feng, C. Jiang, G. Lim, J. Cimini, L.J., G. Feng, and G. Li, "A survey of energy-efficient wireless communications," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 167–178, 2013.

[5] G. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Communications*, vol. 18, no. 6, pp. 28–35, 2011.

[6] Y. Huang, Q. Li, W.-K. Ma, and S. Zhang, "Robust multicast beamforming for spectrum sharing-based cognitive radios," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 527–533, Jan 2012.

[7] L. Fu, Y. Zhang, and J. Huang, "Energy efficient transmissions in MIMO cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 11, pp. 2420–2431, November 2013.

[8] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 79–89, 2007.

[9] C. Shen and M. Fitz, "Dynamic spatial spectrum access with opportunistic orthogonalization," in *43rd Annual Conference on Information Sciences and Systems, 2009.*, 2009, pp. 600–605.

[10] S. Srinivasa and S. Jafar, "Soft sensing and optimal power control for cognitive radio," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3638 –3649, december 2010.

[11] A. Alabbasi, Z. Rezki, and B. Shihada, "Energy efficient scheme for cognitive radios utilizing soft sensing," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2014, pp. 1–6.

[12] D. Feng, C. Jiang, and G. Lim, "A survey of energy-efficient wireless communications," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–12, Feb. 2012.

[13] R. Fischer and J. Huber, "A new loading algorithm for discrete multitone transmission," in *Global Telecommunications Conference, 1996. GLOBECOM '96. 'Communications: The Key to Global Prosperity*, vol. 1, 1996, pp. 724 –728 vol.1.

[14] A. Leke and J. Cioffi, "A maximum rate loading algorithm for discrete multitone modulation systems," in *Global Telecommunications Conference, 1997. GLOBECOM '97., IEEE*, vol. 3, nov 1997, pp. 1514 –1518 vol.3.

[15] A. Lozano, A. Tulino, and S. Verdu, "Optimum power allocation for parallel gaussian channels with arbitrary input distributions," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 3033 – 3051, july 2006.

[16] R. S. Prabhu and B. Daneshrad, "An energy-efficient water-filling algorithm for OFDM systems," in *IEEE International Conference on Communications (ICC), 2010*, 2010, pp. 1–5.

[17] R. Prabhu and B. Daneshrad, "Energy minimization of a QAM system with fading," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4837 –4842, december 2008.

[18] Z. Shi, K. C. Teh, and K. H. Li, "Joint design of sensing and transmission in energy-efficient cognitive radio systems over fading channels," *Communications, IET*, vol. 7, no. 6, pp. 577–584, April 2013.

[19] H. Volos and R. Buehrer, "Cognitive engine design for link adaptation: An application to multi-antenna systems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 9, pp. 2902 –2913, september 2010.

[20] B. A. Fette and B. Fette, *Cognitive Radio Technology (Communications Engineering)*. Newnes, 2006.

[21] G. Miao, N. Himayat, and G. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 545–554, February 2010.

[22] C. Isheden, Z. Chong, E. A. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter."

[23] F. F. Digham, M. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 21–24, Jan. 2007.

[24] A. Cambini and L. Martein, *Generalized Convexity and Optimization: Theory and Applications*. Springer, 2008.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*, 2004.

[26] V. Asghari and S. Aissa, "Resource sharing in cognitive radio systems: Outage capacity and power allocation under soft sensing," in *IEEE Global Telecommunications Conference, GLOBECOM 2008*, 2008, pp. 1 –5.

[27] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert-W function," in *Advances in Computational Mathematics*, 1996, pp. 329–359.

[28] I. R. . Radio Frequecy Micro Devices. (2013, Jun.) Website. [Online]. Available: http://www.rfmd.com/

[29] S. A. Vorobyov, S. Cui, Y. C. Eldar, W. KinMa, and W. Utschick, "Optimization techniques in wireless communications," *EURASIP Journal on Wireless Communications and Networking*, September 2009.

[30] T. C. Clancy, "Formalizing the interference temperature model," *Wireless Communications and Mobile Computing*, vol. 7, no. 9, pp. 1077–1086, 2007. [Online]. Available: http://dx.doi.org/10.1002/wcm.482

[31] "IEEE standard for information technology–telecommunications and information exchange between systems–local and metropolitan area networks–specific requirements part 22.1: Standard to enhance harmful interference protection for low-power licensed devices operating in tv broadcast bands," *IEEE Std 802.22.1-2010*, pp. 1–145, 2010.

[32] G. Arfken and H. J. Weber, *Mmathematical Methods for Physicists*, 6th ed. Elsevier Academic Press, 2005.

[33] S. V. Fomin and R. A. Silverman, *Calculus of Variations*, R. A. Silverman, Ed. Prentice-Hall, Inc., 2000.