

Optimal Cross-Layer Design for Energy Efficient D2D Sharing Systems

Abdulrahman Alabbasi and Basem Shihada

Abstract

In this work, we propose a cross-layer design, which optimizes the energy efficiency of a potential future 5G spectrum-sharing environment, in two sharing scenarios. In the first scenario, underlying sharing is considered. We propose and minimize a modified energy per good bit (MEPG) metric, with respect to the spectrum sharing user's transmission power and media access frame length. The cellular users, legacy users, are protected by an outage probability constraint. To optimize the non-convex targeted problem, we utilize the generalized convexity theory and verify the problem's strictly pseudo-convex structure. We also derive analytical expressions of the optimal resources. In the second scenario, we minimize a generalized MEPG function while considering a probabilistic activity of cellular users and its impact on the MEPG performance of the spectrum sharing users. Finally, derive the associated optimal resource allocation of this problem. Selected numerical results show the improvement of the proposed system compared to other systems.

I. INTRODUCTION

The interest in improving spectrum scarcity and greenness for future wireless systems, including 5G, has been dramatically increasing [1], [2]. The focus on improving the system greenness did not stop at the end user level but also considered the greenness of the backhaul for ultra-dense 5G networks [3]. Furthermore, several energy efficiency (EE) trade-offs have been analyzed, such as, EE versus deployment efficiency metric and EE versus spectrum efficiency [4], [5]. Researchers have addressed several EE metrics through different factors based on different communication layers, e.g., routing, physical (PHY), media access control (MAC) layers, etc., [6], [7], [8]. On the other hand, spectrum scarcity problem can be addressed through several concepts, such as, millimeter wave communication and spectrum-sharing [9]. One of the technologies that might adopt spectrum-sharing in future cellular system is to allow users to

communicate using Device-to-Device (DD) concepts, while other nodes communicate via legacy basestation (BS) concept. In this work, we tackle the improvement of the energy efficiency of spectrum sharing users, called DD user (DDU), while protecting the communication of legacy cellular users, called BS associated user (BSU).

In this context, several work addressed the minimization of the energy efficiency of spectrum-sharing systems from a single layer perspective. The authors of [10] minimized the power to capacity ratio of a multi-carrier cognitive radio (CR) system, while protecting BSU via an average interference constraint. Their optimization problem addresses the minimization of a power to capacity ratio, as an objective function, subject to average CR constraint and quality of service constraint. Their optimization variables include both transmission power and sensing threshold. In [11], authors studied the tradeoff between the energy efficiency and the spectral efficiency. This study addressed the fairness issue among different orthogonal frequency-division multiple access (OFDMA) users. The corresponding near-optimal power allocation of the targeted tradeoff has been derived. Under multiple-input and multiple-output (MIMO) environment, authors of [12] formulated the EE problem as a minimization of power constrained by minimum rate constraint. They derived the optimal beamforming vector and the associated time slot in a time division multiple access (TDMA) network. [As massive MIMO is considered as a solution for 5G networks, authors of \[13\] formulated an energy/cost efficiency problem that addresses wireless communication systems with large number of antennas and radio frequency \(RF\) chains. They have also proposed several algorithms to solve the associated non-concave problem.](#) Authors of [14] have maximized a bounding problem for the energy efficiency while considering the transmission control protocol (TCP) traffic. They have utilized finite state Markov channel and the restless bandit programming techniques to solve the addressed problem. The authors of [15] studied the impact of malicious users on the decision accuracy of cooperative sensing and the associated energy efficiency. They have shown that the fusion rule of the cooperative sensing schemes and the number of malicious nodes hugely impact the system performance. In order to reduce the impact of malicious nodes on energy efficiency, the authors have proposed a low-overhead cryptographic mechanism.

In the same context, several cross-layer design schemes (utilizing several layers' parameters) have been proposed to improve the energy efficiency of different systems. The authors in [16] proposed S-MAC to reduce energy consumption in wireless sensor networks. S-MAC enables nodes to operate on a low-duty-cycle, introducing sleeping time to reserve energy. It also re-introduces the message-passing concept to save energy through controlling the overhead. In [17], authors assumed a quasi-static Rayleigh fading channel, i.e., channel does not change during re-transmission, but it may change independently for new transmissions. Considering a hybrid automatic repeat request (H-ARQ) protocol, authors of [17] devised an optimal power allocation that is obtained recursively for each of the retransmission rounds. Authors of [18] addressed the energy efficiency metric from a cross-layer design perspective. They minimized the transmission energy and normalized transmission energy, where the normalization parameter is a function of retransmission factor, frame length, modulation index and channel coding. In a wireless sensor network environment, authors of [19] utilized the cognitive radio's dynamic spectrum access techniques to enhance the energy consumption of the sensor networks. They considered a cluster-based sensor nodes with inter-and-intra clusters communication periods. They proposed a dynamic channel access technique for intra-cluster communication to reduce the associated energy consumption. They also proposed a joint power allocation and channel access algorithm for improving the energy consumption of inter-cluster communication. Considering a centric based CR network, authors of [20] have considered a scheduling perspective to formulate an energy efficiency problem. Their scheduler considers the diversity among different CR nodes queues and channel capacity. The solution of this problem is obtained via heuristic algorithm. They have also considered an energy consumption minimization problem and energy constraint throughout minimization problem. The authors of [21] analyzed the spectral-energy efficiency trade-off from link and system level perspectives. At the link level, they analyzed the required energy to achieve a certain spectral efficiency while protecting the primary receiver. At the system level, they utilized the extreme-value theorem to derive spectral and energy efficiency.

In this work, we formulate then minimize the average total transmission power (including the cost of retransmissions) to capacity ratio, the introduced metric is called modified energy per

goodbit (MEPG), of the DD system (DDS). The motive behind using MEPG is to investigate the energy cost resulted from employing retransmission protocols, e.g., automatic repeat request (ARQ), and its impact on the total energy efficiency of the system [22]. Another motivation for using MEPG, is to capture the impact of frame-length adaptation, similar concept to frame aggregation [23], on the energy efficiency of the system. We minimize MEPG under two spectrum-sharing scenarios. The first scenario investigates the DDU's MEPG performance while BSU is always active, which is interpreted as underlying sharing. Part of the first scenario has been tackled in our work in [24]. The second scenario considers the impact of a probabilistic model of BSU's activity on DDU's MEPG, while combining underlying and opportunistic sharing schemes. This is a general scenario that tackles more variety of aspects compared to the first scenario. Since both scenarios operate under spectrum-sharing environment, where DDU is assumed to have lower priority than BSU, the DDU must protect BSU transmission by enforcing its outage probability to be lower than a certain threshold. This is called an inband DDS, where the users operating in DD mode and users operate in BS/cellular mode are interfering with each other [25]. After verifying a unique structure of the targeted MEPG problems, we provided an analytical solution along side an iterative algorithm to globally solve the problem.

In this work, our contributions are folded as follows. First, we capture the impact of H-ARQ (by capturing the average number of retransmissions) on the minimum power to Shannon capacity problem, while optimizing with respect to both the transmission power and frame length. Second, the addressed problem is modeled under two sharing scenarios, underlying sharing scenario and a general combined sharing scenario. Third, for both problems, we have verified a unique structure of the objective function, i.e., strict pseudo-convex structure. Fourth, using this structure, analytical expressions of the optimal power and frame length are derived, under both scenarios. These analytical solutions are devised via several optimization techniques, i.e., decomposition and transformation of the variables and the original problem. Fifth, we also propose an alternating iterative algorithm, which uses the derived expressions to guarantee the global minimizer. Several numerical results have been devised to evaluate both scenarios and their impacts on the systems performance, in comparison to: other systems that do not involve

TABLE I
MODEL'S NOTATIONS

Symbol/Operation	Descriptions	Symbol/Operation	Descriptions
P_{pk}	Peak power threshold	$a < b \iff c < d$	The inequality of the left side implies that on the right side and vice versa.
P^{st}	A combination of peak and interference power constraints	ϵ	BSU interference tolerance threshold
P_b	BSU's transmission power	P_d	DDU's transmission power
R_s	DDU's fixed rate	R_b	BSU's fixed rate
ϵ	BSU's outage tolerance threshold	α_1 & α_0	Probability of BSU active & idle
T_s and T_c	Sensing and Communication time	T	$T = T_s + T_c$
γ_d, γ_b	DDU, BSU channel gain	k_c and k_t	Circuit and amplifier constant powers
γ_{db}, γ_{bd}	DT-to-BR, BT-to-DR cross channel gain	KKT	Karush-Kuhn-Tucker
$[G]_b^a$	$\min(\max(G, b), a)$	AWGN	additive white Gaussian noise
L and L_0	Payload frame length and overhead length	i.i.d.	Independent and identically distributed
BSU	BS associated user	EPG	Energy per goodbit
DDU	DD user	MEPG	Modified energy per goodbit
DT	DDU transmitter	CDF	Cumulative density function
DR	DDU receiver	CR	Cognitive radio
BT	BSU transmitter	CSI	Channel state information
BR	BSU receiver	QoS	Quality of service
PDF	Probability density function	SINR	Signal-to-interference-plus-noise ratio

H-ARQ impact, and other sharing technique, such as, opportunistic sharing.

Unlike the work in [18], we consider MEPG metric, i.e., average power, considering average number of retransmissions, divided by the Shannon capacity as the targeted energy efficiency measure. In addition to considering Shannon capacity, our analysis includes the retransmission parameters, which depend on the modulation index, channel coding scheme, packet and overhead length. Furthermore, two different spectrum-sharing scenarios (which suited for 5G systems) have been studied inhere, which is not the case for the work of [18].

The organization of the paper is as follows. Section II describes our system model. Section III discusses the problem formulation of the underlying sharing scenario. Section IV solves the minimization problem of MEPG and finds the optimal solutions. Section V investigates and analyzes the impact of probabilistic BSU model on DDU MEPG's performance. Finally, selected numerical results are presented in section VI.

II. SYSTEM MODEL

In this work, we utilize a general framework, which can be applied to several spectrum-sharing environments. We assume an application of 5G system with two types of nodes/users, as shown in Fig. 1. The first type, are users associated with the BSs, called BSUs. The second type, are users operating in a DD communication mode, called DDUs. Since our main objective is to

address the energy efficiency of spectrum-sharing systems, we link the DDS and BS associated system (BSS) to sharing concepts as follows. We consider that DDUs have lower priority to access the spectrum in compared to the BSUs. Therefore, we enforce a constraint on the DDU's transmission to protect the BSU from the DDU's interference. **In this work, we consider the inband underlying DDS [25]. That is, the users in the same mode are operating in different bands such that they do not interfere with each other. Whereas, they interfere with the users from the other mode, i.e., DDUs interfere with BSUs and vice versa.** We also assume that BS allocates similar size of resource blocks, i.e., frequency band, to all users. BS is allowed to allocate more resource blocks to the BSUs compared to the DDUs. This assumption guarantees that each DDU interferes with a single BSU over the allocated block. Herein, we denote the channel gain between DDU transmitter (DT) and DDU receiver (DR) as γ_d and the channel gain between BSU transmitter (BT) and BSU receiver (BR) as γ_b . Whereas, the cross-link channel gain between DT and BR is denoted as γ_{db} and the cross-link channel gain between BT and DR is denoted as γ_{bd} . The corresponding means of the channels gains are μ_d , μ_b , μ_{db} , μ_{bd} , respectively, whereas, the variances are expressed as σ_d^2 , σ_b^2 , σ_{db}^2 , σ_{bd}^2 , respectively. All channels are assumed to be independent and identically distributed (i.i.d.). Through a feedback channel, we assume that DDU knows the instantaneous γ_d , whereas, γ_b , γ_{bd} , and γ_{db} are not available at DDU's side. We assume that BS shares the statistical information about the cross-links channels gains between DDU and BSU, i.e., σ_{db}^2 and σ_{bd}^2 . This assumption is practical since the exchange of this information does not impose a huge overload on the control channel. Because it does not require an instantaneous update on channel state information (CSI) of the cross-link gains.

Several assumptions are utilized to formulate our targeted problems. We assume that BSU is transmitting with a maximum power policy, i.e., P_b . This is justified by the fact that we are addressing the worst case of BT interference toward DR. On the other hand, DT has an adaptive power policy. We also assume that DDU decodes the BSU interference as noise, since it does not know γ_{bd} , hence, it is unable to perform any interference mitigation technique. The proposed system considers a framework that takes into consideration minimizing the DDU's MEPCG with respect to two variables (DDU's transmission power and DDU's MAC layer's

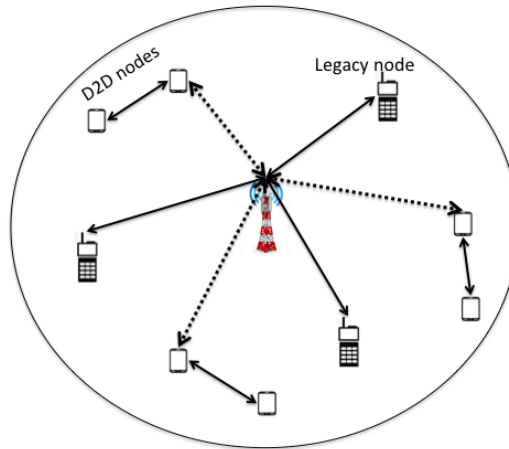


Fig. 1. Spectrum sharing framework with BSUs and D2Ds operating in the same cell.

frame length). Hence, the problem's degrees of freedom are increased. We include two spectrum-sharing scenarios in our proposed minimization problem. The first scenario considers underlying sharing, without sensing information, whereas, the second scenario considers combined sharing and requires sensing information. The second scenario considers a probabilistic activity of BSU, which can be modeled via a bernoulli random variable, i.e., combined underlying and opportunistic sharing. The information about the activity of BSU can be obtained via two approaches. The first approach, which is considered in this work, requires the DDU to sense the spectrum. The other approach requires the BS to send this information to the DDU, but it results in a considerable overhead on the control channel, thus this approach is not considered in here. Note that we do not consider the BS offloading improvement results from using DDS.

III. PROBLEM FORMULATION & STRUCTURE OF UNDERLAYING SCENARIO

In this section, we present and investigate the problem formulation of the proposed system in underlying sharing scenario. In the first sub-section, we introduce the targeted optimization problem and the necessary assumptions to solve it. In the second sub-section, we investigate the problem structure, by verifying that the problem is not convex, but it follows a strict pseudo-convexity structure, which is going to be utilized later in the solution section.

A. Problem Formulation

The proposed problem addresses the minimization of MEPG of the DDS. The objective function, MEPG, is expressed in (1a) which contains an average number of retransmissions, \mathfrak{R}_t , multiplied by each transmission power cost, $\frac{L_F}{L} k_t P_d + k_c$, divided by the normalized capacity, $\frac{L}{L_F} \log \left(1 + \frac{P_d \gamma_d}{1 + P_b \sigma_{bd}^2} \right)$. The minimization of MEPG is subject to three constraints, as follows. The first constraint, \mathfrak{C}_1 in (1b), reserves the BSU's outage probability, $\Pr \left[\mathcal{I}_b < R_b \middle| \gamma_d \right]$, to be lower than a threshold ϵ . The peak power constraint, \mathfrak{C}_2 , is expressed in (1c). The third constraint, \mathfrak{C}_3 in (1d), guarantees that DDU achieves a minimum throughput of R_s . The proposed problem formulation is expressed as follows,

$$\min_{P_d, L} \quad \mathcal{E}_e(P_d, L) = \frac{\mathfrak{R}_t \left[\frac{L_F}{L} k_t P_d + k_c \right]}{\frac{L}{L_F} \log \left(1 + \frac{P_d \gamma_d}{1 + P_b \sigma_{bd}^2} \right)} \quad (1a)$$

$$\text{s.t.} \quad \mathfrak{C}_1 : \Pr \left[\mathcal{I}_b < R_b \middle| \gamma_d \right] \leq \epsilon. \quad (1b)$$

$$\mathfrak{C}_2 : P_d \leq P_{pk} \quad (1c)$$

$$\mathfrak{C}_3 : \log \left(1 + \frac{P_d \gamma_d}{1 + P_b \sigma_{bd}^2} \right) \geq R_s, \quad (1d)$$

where k_c and k_t are assigned parameters which are associated with the circuit power of the radio devices and power amplifier constant power consumption. **The minimum rate threshold is R_s .** The MAC layer frame length $L_F = L + L_0$ includes the information bits, i.e., L , and the overhead bits, i.e., L_0 . The DDU's and BSU's transmission powers are designated as P_d and P_b , respectively. The numerator of (1a) presents the DDU's average transmission power times the average number of retransmissions. Whereas the denominator presents a normalized capacity of DDU. Since we assume that DDU does not have knowledge of γ_{bd} , the targeted mutual information of DDU is derived along similar lines as in [26], i.e., $\mathcal{I}_s = \log \left(1 + \frac{P_d \gamma_d}{1 + P_b \sigma_{bd}^2} \right)$.¹ Given that the DDU's desired signal is X_d and received signal is Y , the DDU's mutual information is derived using the following expression $\mathcal{I}(Y; X_d | h_d) = h(X_d | h_d) - h(X_d | Y, h_d)$. The first term is obtained given that the input signal has a Gaussian entropy, i.e., $h(X_d | h_d) = \log(2\pi e P_d)$. The second term is obtained using the minimum linear mean square-error estimator (LMSE), such that, $h(X_d | Y, h_d) \leq \log \left(2\pi e \frac{P_d P_b \sigma_{bd}^2 + P_d}{\gamma_d P_d + \sigma_{bd}^2 P_b + 1} \right)$. The detailed proof of this derivation is mentioned in Appendix A. It follows that the DDU's normalized capacity is then obtained as

¹The parameter $1 + P_b \sigma_{bd}^2$ is obtained via sensing BSU's signal, i.e., $\mathbb{E}\{y_{bd} y_{bd}^*\} = 1 + P_b \sigma_{bd}^2$, where y_{bd} is BSU's signal at DDU. Both σ_{bd}^2 and σ_{db}^2 are made available at DDU's side via blind sensing or CR centralized engine (BS).

$\frac{L}{L_F} \log \left(1 + \frac{P_d \gamma_d}{1 + P_b \sigma_{bd}^2} \right)$. The parameter ϵ is a tolerance threshold for the violation of BSU targeted outage probability. Let us denote $P_I = 1 + P_b \sigma_{bd}^2$. Similar to DDU case, the targeted BSU's mutual information is expressed as $\mathcal{I}_b = \log \left(1 + \frac{P_b \gamma_b}{1 + P_d \sigma_{db}^2} \right)$. Note that \mathfrak{R}_t is the average number of frame retransmissions, expressed as, $\mathfrak{R}_t = \frac{1}{1 - \mathfrak{F}_{er}}$, where \mathfrak{F}_{er} is the frame error rate, which is approximated under a Rayleigh fading channel as in [27],

$$\mathfrak{F}_{er} \simeq 1 - \exp \left(-\frac{\gamma_w}{\zeta_{db}} \right), \quad (2)$$

where $\zeta_{db} = P_d / (1 + P_b \sigma_{bd}^2)$. The threshold γ_w is approximated as, $\gamma_w \simeq k_M \log(L_F) + b_M$, where k_M and b_M are the related parameters to the coding and modulation schemes [18]. Finally, the number of retransmissions is rewritten as,

$$\mathfrak{R}_t \simeq \exp \left(\frac{k_M \log(L_F) + b_M}{\zeta_{db}} \right). \quad (3)$$

It is clear that both constraints \mathfrak{C}_1 and \mathfrak{C}_2 can be combined, whereas constraint \mathfrak{C}_3 can be rewritten as a lower bound constraint, thus (1) is rewritten as follows,

$$\min_{P_d, L} \mathcal{E}_e(P_d, L) = \frac{\mathfrak{R}_t \frac{L_F}{L} \left[\frac{L_F}{L} k_t P_d + k_c \right]}{\log \left(1 + \frac{P_d \gamma_d}{P_I} \right)} \quad (4a)$$

$$\text{s.t.} \quad P_d \leq \min \{ Q_I, P_{pk} \} = P^{st} \quad (4b)$$

$$P_d \geq \frac{[\exp(R_s) - 1] P_I}{\gamma_d} = P_{mn}. \quad (4c)$$

Note that the capacity normalization parameter $\frac{L}{L_F}$ is moved to the numerator of (4a). Constraints \mathfrak{C}_1 and \mathfrak{C}_2 are converted into a short-term power constraint, i.e., (4b). The short-term power constraint Q_I is obtained from \mathfrak{C}_1 as follows,

$$\Pr \left[\log \left(1 + \frac{P_b \gamma_b}{P_d \sigma_{db}^2 + 1} \right) < R_b \middle| \gamma_d \right] \leq \epsilon \implies F_{\gamma_b | \gamma_d} \left(\gamma_b \leq \frac{(P_d (\gamma_s) \sigma_{db}^2 + 1) (e^{R_b} - 1)}{P_b} \right) \leq \epsilon \quad (5a)$$

$$\implies P_d \leq \left[\frac{F_{\gamma_b | \gamma_d}^{-1}(\epsilon) P_b}{(e^{R_b} - 1) \sigma_{db}^2} - \frac{1}{\sigma_{db}^2} \right]^+ = Q_I, \quad (5b)$$

where (5a) is obtained from the independence between γ_b and γ_d and (5b) is obtained from the fact that $F_{\gamma_b | \gamma_d}$, being a cumulative density function (CDF), is a monotonically non-decreasing function. Note that BSU knows the channel gain γ_b , however, this constraint is enforced at DDU's side. Therefore, we condition on γ_d and not on γ_b , which has been replaced by its distribution

structure, i.e., $F_{\gamma_b|\gamma_d}^{-1}(\cdot)$, after solving the probability.

B. Problem Structure

In order to analyze the objective function, MEPG, we begin by verifying that it is not convex for all values of P_d and L . Rewriting the numerator of $\mathcal{E}_e(P_d, L)$ as follows,

$$\begin{aligned}\mathcal{E}_n(P_d, L) &= e^{\left(\frac{a'(L)}{P_d}\right) \frac{L_F}{L}} \left[\frac{L_F}{L} k_t P_d + k_c \right] \\ &= \mathcal{E}_n^{(1)}(P_d, L) + \mathcal{E}_n^{(2)}(P_d, L)\end{aligned}\quad (6)$$

where $a'(L) = k'_M \log(L_F) + b'_M$, $k'_M = k_M P_I$ and $b'_M = b_M P_I$. We define the two terms of (6) as follows, $\mathcal{E}_n^{(1)}(P_d, L) = \exp\left(\frac{a'(L)}{P_d}\right) \left(\frac{L_F}{L}\right)^2 k_t P_d$ and $\mathcal{E}_n^{(2)}(P_d, L) = \exp\left(\frac{a'(L)}{P_d}\right) \frac{L_F}{L} k_c$. Note that, hereafter, we use the notations $\mathcal{E}_n(P_d, L)$, $\mathcal{E}_n^{(1)}(P_d, L)$ and $\mathcal{E}_n^{(2)}(P_d, L)$ interchangeably with \mathcal{E}_n , $\mathcal{E}_n^{(1)}$ and $\mathcal{E}_n^{(2)}$. To check the convexity of MEPG (with respect to L), we test the convexity of \mathcal{E}_n . After analyzing the structure of $\mathcal{E}_n^{(2)}$, we conclude that the convexity of $\mathcal{E}_n^{(2)}$ with respect to L depends on the value of the allocated power (unlike the results in [18]), as follows,

$$\forall \frac{k'_M}{P_d} > 1 \implies \begin{aligned} \mathcal{E}_n^{(2)} \text{ is convex if: } & \frac{L_0 P_d}{k'_M - P_d} > L \\ \mathcal{E}_n^{(2)} \text{ is concave if: } & \frac{L_0 P_d}{k'_M - P_d} < L \end{aligned}\quad (7a)$$

$$\forall \frac{k'_M}{P_d} < 1 \implies \begin{aligned} \mathcal{E}_n^{(2)} \text{ is convex if: } & \frac{L_0 P_d}{k'_M - P_d} < L \\ \mathcal{E}_n^{(2)} \text{ is concave if: } & \frac{L_0 P_d}{k'_M - P_d} > L. \end{aligned}\quad (7b)$$

The results in (7) verifies that $\mathcal{E}_n^{(2)}$ is not a convex function in L . Following a similar line it is easy to show that $\mathcal{E}_n^{(1)}$ is also not convex in L . It is also noted that \mathcal{E}_e is a fractional non-convex function of P_d . Therefore, it is necessary to find an alternative structure of \mathcal{E}_e , with respect to P_d and L , to guarantee the global optimal solution. In [28], it is proven that to find the global optimal solution, it is sufficient to prove the pseudo-convexity structure of the problem, with respect to its variables, and satisfy the Karush-Kuhn-Tucker (KKT) conditions. Therefore, we verify the strict pseudo-convexity structure of \mathcal{E}_e with respect to each of the optimization variables, i.e., P_d and L , respectively².

²However, note that verifying the pseudo-convexity structure, separately, with respect to each variable does not result in a jointly global optimal solution. Therefore, we verify the strictness of the function. This strict pseudo-convex structure enables us to utilize the results in [29]. Hence, we utilize an alternating algorithm to iterate over each optimal solution (optimal with respect to each variable) until it reaches a global joint optimal solution [29].

Lemma 1. *The numerator of $\mathcal{E}_e(P_d, L)$ is strictly convex with respect to P_d .*

Proof. It is clear that $\mathcal{E}_n^{(2)}$ is a strictly convex function with respect to P_d . To prove the strict convexity property of $\mathcal{E}_n^{(1)}$ we devise the second order derivative as follows,

$$\frac{\partial^2 \mathcal{E}_n}{\partial P_d^2} = e^{\left(\frac{a'(L)}{P_d}\right)} \frac{k_t L_F}{L P_d^3} [-P_d + a^2(L) + a'(L)P_d] > 0. \quad (8)$$

It is enough to recall that $a'(L) > 1$ to verify (8). Since the sum of two strictly convex function, $\mathcal{E}_n^{(1)}$ and $\mathcal{E}_n^{(2)}$, results in strictly convex function. Thus, \mathcal{E}_n is strictly convex function in P_d . \square

It is observed that $\log\left(1 + \frac{P_d \gamma_d}{P_I}\right)$ is strictly concave with respect to P_d . In the following proposition, to investigate the structure of \mathcal{E}_e with respect to P_d , we utilize the fact that the numerator and denominator of \mathcal{E}_e are strictly convex and strictly concave, respectively.

Proposition 1. *The objective function of problem (4), $\mathcal{E}_e(P_d)$, is strictly pseudo-convex in P_d .*

Proof. To prove proposition 1, we show that the ratio between a strictly convex function and a strictly concave function results in a strictly pseudo-convex function. To show this, two conditions must be satisfied [28]:

- 1) The function $\mathcal{E}_e(P_d)$ is a strictly quasi-convex function.
- 2) There exist a local minimum P_d^* , i.e., $\nabla \mathcal{E}_e(P_d) = 0$.

A strictly quasi-convex function is defined as follows, $\mathcal{E}_e\left(\lambda P_d^{(1)} + (1 - \lambda)P_d^{(2)}\right) < \max\left\{\mathcal{E}_e(P_d^{(1)}), \mathcal{E}_e(P_d^{(2)})\right\}$. For the sake of analysis, we rewrite $\mathcal{E}_e(P_d)$ as a ratio between a numerator and a denominator terms as follows, $\mathcal{E}_e(P_d) = \frac{\mathcal{E}_n(P_d)}{\mathcal{E}_d(P_d)} = \frac{\Re_t \frac{L_F}{L} \left[\frac{L_F}{L} k_t P_d + k_c\right]}{\log\left(1 + \frac{P_d \gamma_d}{P_I}\right)}$. Strict convexity and concavity properties of each $\mathcal{E}_n(P_d)$ and $\mathcal{E}_d(P_d)$ are utilized to show the strict quasi-convexity of $\mathcal{E}_e(P_d)$ as follows,

$$\mathcal{E}_e\left(\lambda P_d^{(1)} + (1 - \lambda)P_d^{(2)}\right) < \frac{\mathcal{E}_n\left(P_d^{(1)}\right)}{\mathcal{E}_d\left(P_d^{(1)}\right)} \left[\mathcal{E}_d\left(\lambda P_d^{(1)} + (1 - \lambda)P_d^{(2)}\right)\right], \quad (9)$$

where we assume that $\mathcal{E}_e\left(P_d^{(2)}\right) < \mathcal{E}_e\left(P_d^{(1)}\right)$, (9) is valid because of the strict concave property of $\mathcal{E}_d(P_d)$. Detail of the above proof can be found in [10].

The second step in the proof is to show that $\nabla \mathcal{E}_e(P_d^*) = 0$. This step is rigorously derived in Appendix B.

By combining (9) and the results from Appendix B, we conclude that \mathcal{E}_e is strictly pseudo-convex. Hence, utilizing the results from [28] the P_d minimizer of problem (1) is obtained by satisfying the KKT conditions of the corresponding problem. □

In the following proposition, we show the structure of problem (4) with respect to frame length, L .

Proposition 2. *The objective function of problem (4) is strictly pseudo-convex in L .*

Proof. The detailed proof is given in Appendix C. □

Using both Proposition 1 and Proposition 2 it is concluded that problem (4) is strictly pseudo-convex with respect to transmission power and frame length. In the next section, we utilize this conclusion, about the problem structure, to derive the optimal solution.

IV. SOLUTION OF UNDERLAYING SCENARIO

In this section, we provide the optimal power policy and frame length that minimizes $\mathcal{E}_e(P_d, L)$. We begin by transforming problem (4) by using parametric/fractional optimization techniques. It is proven in [30] that utilizing fractional optimization the minimization a function of a similar structure to (4a) is equivalent to minimizing the following function,

$$g(P_d, L, t) = e^{\left(\frac{a'(L)}{P_d^n}\right) \frac{L_F}{L}} \left[\frac{L_F}{L} k_t P_d + k_c \right] - t \log\left(1 + \frac{P_d \gamma_d}{P_I}\right), \quad (10)$$

with respect to t , P_d , and L . It follows that the minimization problem in (4) is equivalent to the following problem,

$$\min_{t, P_d, L} g(P_d, L, t) \quad \text{s.t.} : P_d \leq P^{st}; \quad \& \quad P_d \geq P_{mn}. \quad (11)$$

Due to the nonlinear function of transmission power ($g(P_d)$), i.e., P_d exists as $\frac{a'(L)}{P_d}$ in the exponential term and inside the logarithm in (10), the expression of the optimal P_d is difficult

to obtain. Therefore, we convert problem (11) to an equivalent problem by introducing a new optimization variable, P_{dn} , and an equality constraint as follows,

$$\min_{t, L, P_{mn} \leq \{P_d, P_{dn}\} \leq P^{st}} e^{\left(\frac{\alpha'(L)}{P_{dn}}\right) \frac{L_F}{L} \left[\frac{L_F}{L} k_t P_d + k_c \right]} - t \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) \quad (12a)$$

$$P_{dn} = P_d. \quad (12b)$$

Introducing the equality constraint, in (12b), is a key enabler of our solution methodology. This is because it separates the P_d term in the exponential and the P_d term in the logarithm.

The corresponding Lagrangian function of problem (12) is expressed as follows,

$$\mathfrak{L} = e^{\frac{\gamma_w P_I}{P_{dn}} \frac{L_F}{L} \left[\frac{L_F}{L} k_t P_{dn} + k_c \right]} - t \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) - \mu [P_{dn} - P_d]. \quad (13)$$

where μ is the Lagrangian multiplier associated with the equality constraint in (12b). The upper-bound constraint on P_d is applied on the optimal power expression as introduced below.

The optimal P_d is expressed as follows,

$$P_d^* = \left[\frac{\lambda t L e^{-\left(\frac{\gamma_w P_I}{P_{dn}}\right)} - \frac{P_I}{\gamma_d}}{\mu L_F} \right]_{P_{mn}}^{P^{st}}. \quad (14)$$

The optimal P_{dn} is expressed as follows,

$$P_{dn}^* = \left[\frac{\gamma_w P_I}{\left(2W \left[\frac{b \gamma_w P_I}{2a} \sqrt{\frac{a}{b}} \right] \right)} \right]_{P_{mn}}^{P^{st}}, \quad (15)$$

where $a = \frac{\lambda t L}{L_F} \log \left(1 + \frac{P_d \gamma_d}{1 + P_b \sigma_{bd}^2} \right)$ and $b = \frac{L_F}{L} k_t \lambda - \mu$. The function $W[\cdot]$ is the Lambert function. Finding both (14) and (15) is a straightforward implementation of generalized convexity theory and satisfying the associated KKT conditions. The optimal value of t is obtained numerically using a bi-sectional algorithm. This guarantees an optimal solution because the problem is quasi-convex. The optimal value of the Lagrangian variable μ is obtained by finding the roots of the following nonlinear function,

$$f_\mu(\mu) = \frac{\gamma_w P_I}{2W \left[\frac{b \gamma_w P_I}{2a} \sqrt{\frac{a}{b}} \right]} - \frac{\lambda t L e^{-\left(\frac{\gamma_w P_I}{P_{dn}}\right)}}{\mu L_F} + \frac{P_I}{\gamma_d} = 0. \quad (16)$$

The optimal frame length is expressed as follows,

$$L^* = \frac{L_0 \frac{P_{dn}}{k_M P_I}}{2(k_c + P_{dn} k_t)} \left[- (k_M P_I k_t - 2P_{dn} k_t - k_c) + \sqrt{(k_M P_I k_t - 2P_{dn} k_t - k_c)^2 + 8k_t k_M P_I (k_c + k_t P_{dn})} \right]. \quad (17)$$

The above optimal expression of the frame length is found by deriving the partial derivative of (13) with respect to L . Then, the zero of the resulted function is reformulated to the exact expression of (17). It is necessary to note that L^* is an increasing function of the transmission power. This is aligned with the experimental results which shows that the frame length increases with the increase of the signal-to-interference-plus-noise ratio (SINR). This reduces the overhead which is generated by adding multiple frame headers to small frames [31].

Finally, in order to guarantee a global joint optimal solution of P_d and L we propose, in Algorithm 1, an alternating optimization algorithm. Because of the strictly quasi-convex structure of the problem with respect to each variable, it is verified that the proposed algorithm results in jointly global minimum value of \mathcal{E}_e [29]. The parameter δ is the stopping criteria of the algorithm

Algorithm 1: Proposed Algorithm

input : $\delta, \epsilon, \alpha, k_t, k_c, k_M, b_M, P_{pk}, \gamma_d, P_I, L_0$
1 Initialize: $P_d^{(0)} = P_{pk}, P_{dn}^{(0)} = P_{pk}, L^{(0)} = L_{max}, cond = True;$
2 $q = 1$
3 while $cond$ **do**
4 Find a feasible value of t using a bisectional algorithm.
5 Find μ by solving the KKT condition associated with (16), given fixed $P_d = P_d^{(q-1)}, P_{dn} = P_{dn}^{(q-1)}$, and $L = L^{(q-1)}$.
6 By finding μ^* we guarantee that $P_d = P_{dn}$. Thus, DDU's power is found as $P_d^{(q)} = P_d^*$, in (14), given $L = L^{(q-1)}$.
7 By substituting μ^* in (15) we find $P_{dn}^{(q)} = P_{dn}^*$, given $L = L^{(q-1)}$.
8 The optimal value of frame length, L^* , is found using (17) given $P_d = P_d^{(q)}, P_{dn} = P_{dn}^{(q)}$.
9 Evaluate $\mathcal{E}_e^{(q)}(P_d^{(q)}, L^{(q)})$ as in (4a).
10 **if** $\left\| \mathcal{E}_e^{(q)} - \mathcal{E}_e^{(q-1)} \right\| < \delta$ **then:** $cond = False$
11 $q = q+1;$
12 end
output: $\{P_d^{(q)}, L^{(q)}\}$

and L_{max} is the maximum frame length. We initialize t using the bisectional algorithm, and initialize P_d, P_{dn} , and L by initial values. It follows that we find both μ as in (16). We then find, in order, $P_d^{(q)}, P_{dn}^{(q)}$, and $L^{(q)}$. Keep iterating and updating the solution until the difference between the current $\mathcal{E}_e^{(q)}$ and the previous one, $\mathcal{E}_e^{(q-1)}$, is within a pre-selected threshold, δ .

V. IMPACT OF BSU'S PROBABILISTIC ACTIVITY (COMBINED SCENARIO)

In this section, we investigate the impact of probabilistic BSU's activity on DDU's energy efficiency performance. We first formulate the targeted problem. Then, we derive the associated optimal DDU's transmission power and frame length.

BSU's activities are captured as a Bernoulli random variable. The impact of BSU's sensed activity on the DDU's performance is captured by reformulating problem (4) to problem (18). The objective function in (18a) differs from that in (4) by averaging over BSU's activity in the retransmission term, i.e., \mathfrak{R}_t at the numerator, and the capacity term, at the denominator, in addition to including the sensing power consumption. We also normalize all the terms in the objective by the transmission time ratio, $\frac{T_c}{T}$, or the sensing time ratio, $\frac{T_s}{T}$. Constraint (18c) differs from the minimum rate constraint of problem (4) by averaging over BSU's activity and normalization by transmission time ratio. The extended MEPG problem is expressed as follows,

$$\min_{P_d, L} \frac{\overbrace{\left[\alpha_1 e^{\left(\frac{a'(L)}{P_d}\right)} + \alpha_0 e^{\left(\frac{a(L)}{P_d}\right)} \right]}^{\mathfrak{R}_t, \text{ averaged over BSU's active/idle probability}} \overbrace{\left[\frac{T_c L_F}{L T} \left[\frac{L_F}{L} k_t P_d + k_c \right] \right]}^{\text{Normalized transmission power}} + \overbrace{\frac{P_{sns} T_s}{T}}^{\text{Normalized sensing power consumption}}}{\underbrace{\frac{T_c}{T} \left[\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d) \right]}_{\text{Normalized DDU capacity, averaged over BSU's activity}}} \quad (18a)$$

$$\text{s.t.} \quad P_d \leq \min \{ Q_I, P_{pk} \} \quad (18b)$$

$$\underbrace{\frac{T_c}{T} \left[\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d) \right]}_{\text{Normalized DDU capacity, averaged over BSU's activity}} \geq R_s, \quad (18c)$$

where α_0 and α_1 are probabilities of BSU being idle and active, respectively, and $a(L) = k_M \log(L_F) + b_M$, recall that $a'(L) = a(L)P_I$. Note that $\alpha_1 = 1 - \alpha_0$. It is also noted that $e^{\left(\frac{a'(L)}{P_d}\right)} > e^{\left(\frac{a(L)}{P_d}\right)}$ and $\log \left(1 + \frac{P_d \gamma_d}{P_I} \right) < \log (1 + P_d \gamma_d)$ because of the interference effect from BT on DR. The parameters T , T_c , and T_s are the whole frame time, communication time, and sensing time, respectively. Given that these periods are related as $T = T_c + T_s$. The impact of the statistical constraint in (1b) is already included in Q_I of (18b). The impact of BSU's probabilistic activity appears in the numerator and denominator of the DDU's energy efficiency metric, in (18a), and it also appears in the minimum quality of service (QoS) constraint (18c). For instance, if BSU is active with high probability, $\alpha_1 \gg \alpha_0$, the influence of $e^{\left(\frac{a'(L)}{P_d}\right)}$ on

increasing DDU's MEPG will be higher than that of $e^{\left(\frac{\alpha(L)}{P_d}\right)}$. High probability of BSU existence will similarly affect the denominator of DDU's MEPG. The addition of constraint (18c) is to guarantee a minimum quality of service for the DDU's rate performance. This constraint also highlights the impact of the BSU's activity on the DDU's EE performance.

Several optimization techniques and steps are followed in order to solve problem (18). We begin by transforming the problem, such that an insightful solution is found, as in the following lemma.

Lemma 2. *Problem (18) is equivalent to the following problem,*

$$\min_{P_d, P_{de}, L, t} \left[\alpha_1 e^{\left(\frac{\alpha(L)}{P_{de}}\right)} + \alpha_0 e^{\left(\frac{\alpha(L)}{P_{de}}\right)} \right] \frac{T_c L_F}{T L} \left[\frac{L_F}{L} k_t P_d + k_c \right] + P_{sns} \frac{T_s}{T} - \frac{t T_c}{T} \left[\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d) \right] \quad (19a)$$

$$s.t. \quad P_d \leq \min \{ Q_I, P_{pk} \} \quad (19b)$$

$$\frac{T_c}{T} \left[\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d) \right] \geq R_s \quad (19c)$$

$$P_d = P_{de}, \quad (19d)$$

where P_{de} is a new optimization variable. Constraint (19d) is included to guarantee the equivalence between both P_d and P_{de} .

Proof. A detailed proof of the transformation steps is given in Appendix D. \square

The main merit of the reformulated problem in Lemma 2 is the reduced complexity of the original problem with respect to P_d .

Hereafter, we address the necessary steps to obtain the optimal value of each variable, i.e., $\{P_d, P_{de}, L, t\}$, utilizing different theorems and techniques. In lines with Sec. III, we begin by addressing the structure of problem (19), we then find the optimal frame length followed by the optimal transmission power. The structure of problem (19) is addressed in the following lemma.

Lemma 3. *Problem (19) is strictly quasi-convex with respect to each of the optimization variables, i.e., $\{P_d, P_{de}, L, t\}$.*

Proof. The detailed proof of Lemma 3 is given in Appendix E.

□

Utilizing the geometric optimization techniques, we are able to find an expression for the optimal frame length, which is provided in the following proposition.

Proposition 3. *The optimal frame length that minimizes problem (19) is expressed as follows,*

$$\hat{L}_F = \frac{\beta' - \beta}{\left[L_0 t \Gamma - \sqrt{L_0^2 t^2 \Gamma^2 + 2(\beta - \beta') t \Gamma} \right]} \quad (20)$$

where $\Gamma = \alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d)$, $\beta' = \frac{k'_M}{P_d}$, and $\beta = \frac{k_M}{P_d}$.

Proof. The detailed proof of Proposition 3 is given in Appendix F. However, a sketch of the proof is given as follows. We first change the structure of the objective function, using fractional programming. We also use the intuitive approximation as $\frac{L_F}{L} k_t P_d \gg k_c$. Then we utilize the geometric optimization techniques, change of variables, and transformation of objective function to find the optimal value of L_F . Finally, we apply the generalized convexity theory, construct the Lagrangian function, and satisfy the KKT condition to obtain the optimal L_F .

□

The optimal transmission power value is found, in the following proposition, through several steps,

- 1) Minimizing problem (19) with respect to each P_d and P_{de} .
- 2) Utilizing the generalized convexity theory, we find the optimal Lagrangian multipliers λ and μ that satisfy the associated KKT conditions with constraints (19c) and (19d), respectively.

Finding the optimal μ^* results in $P_d^* = P_{de}^*$.

Proposition 4. *The optimal value of P_d which minimizes (19) is expressed as follows,*

$$\hat{P}_d = \left[\frac{-\left[\frac{A}{B}(1 + P_I) - 1 \right] + \sqrt{\left(\left[\frac{A}{B}(1 + P_I) - 1 \right]^2 - 4 \frac{A}{B} \left[P_I \left(\frac{A}{B} - \alpha_0 \right) - \alpha_1 \right] \right)}}{2 \frac{A}{B} \gamma_d} \right]_0^{P^{st}}, \quad (21)$$

where $A = \left[\alpha_1 e^{\left(\frac{a'(L)}{P_{de}} \right)} + \alpha_0 e^{\left(\frac{a(L)}{P_{de}} \right)} \right] \left(\frac{L_F}{L} \right)^2 \frac{T_c}{T} k_t - \mu$ and $B = \frac{T_c}{T} (t - \lambda) \gamma_d$.

The optimal value of P_{de} is obtained as follows,

$$\hat{P}_{de} = \left[\sqrt{\frac{a(L) + a'(L)}{\mu}} \right]^{P^{st}}, \quad (22)$$

The Lagrangian multiplier μ which is associated with constraint (19d) is obtained by solving the zeros of the following function,

$$f(\mu) = -\sqrt{\frac{a(L) + a'(L)}{\mu}} - \frac{[\frac{A}{B}(1 + P_I) - 1]}{2^{\frac{A}{B}}\gamma_d} + \frac{\sqrt{([\frac{A}{B}(1 + P_I) - 1])^2 - 4^{\frac{A}{B}} [P_I (\frac{A}{B} - \alpha_0) - \alpha_1]}}{2^{\frac{A}{B}}\gamma_d} \quad (23)$$

The Lagrangian multiplier λ is found by solving the dual problem and the KKT condition associated with constraint (19c), i.e., finding the zeros of the following function,

$$f(\lambda) = \alpha_1 \log \left(1 + \frac{\hat{P}_d(\lambda)\gamma_d}{P_I} \right) + \alpha_0 \log \left(1 + \hat{P}_d(\lambda)\gamma_d \right) - \frac{TR_s}{T_c} \quad (24)$$

note that the expression of $\hat{P}_d(\lambda)$ is similar to that of \hat{P}_d , which is found in (21).

Proof. To verify Proposition 4, we utilize several steps as follows. We first construct the Lagrangian function to solve for the optimal P_d , via means of satisfying KKT conditions. We then divide the problem into two sub-problems with respect to P_{de} , two new variables and equality constraints are introduced, to construct a tractable solution. We also utilize the theory of objective transformation and derive the optimal value of P_{de} . The detailed proof of Proposition 4 is given in Appendix G. □

The optimal power expression in (21) does not lead to a straightforward insight about how the optimal power changes with the change of the frame length. Therefore, In Fig. 2, we plot the optimal power as a function of frame length. We observe that the optimal transmission power, which requires convergence to the optimal μ^* , is a quasi-concave function of the frame length. It is complex to analytically derive this property, since the optimal transmission power is a function of the optimal μ^* , which is found by using iterative algorithm. Inhere; we utilized the trust region algorithm to solve (23) for the optimal μ^* , which is used in plotting Fig. 2. It is also shown in Fig. 2, that both P_d and P_{de} are equivalent, which is guaranteed by finding the optimal μ .

The residual error resulted from solving (23), using the trust region algorithm, is in the order of $< 10^{-11}$. We have also found that the number of iteration required to solve for optimal μ is not large, i.e., ≤ 13 . We have tuned several system parameters to verify these observations, and we found that these observations, small residual error and small number of iterations required for the convergence, are always valid, given that we are operating in the feasibility region.

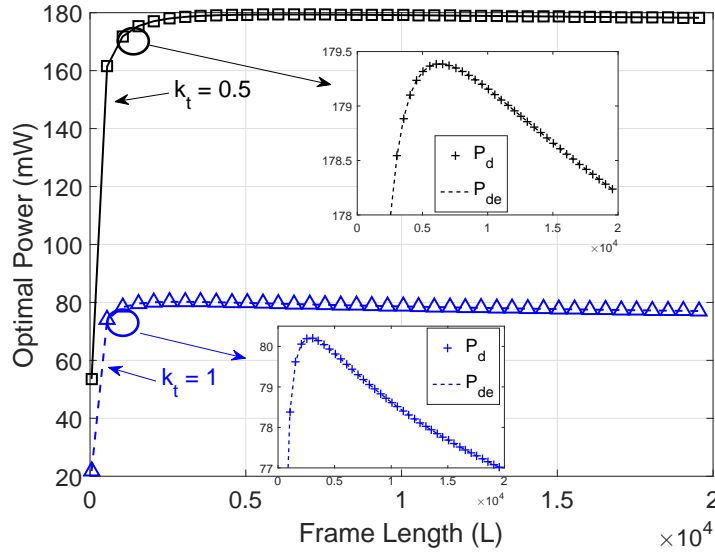


Fig. 2. Optimal power versus frame length, via finding the optimal μ

In similar lines with Alg. 1 and utilizing the results in Lemma 3, Proposition 3, and Proposition 4, it is straightforward to construct an algorithm which guarantees the joint global optimal solution using the previously derive expressions, in (20), (21), and (23).

VI. NUMERICAL EVALUATION

In this section, we evaluate DDU's MPEG performance while changing the system's parameters via numerical methods. Note that in most of the evaluation results we use energy per goodbit (EPG) notation instead of MPEG, for generalization issues. The impact of changing the system parameters on the DDU's energy efficiency performance is studied. We divide the numerical evaluations into two parts. The first part addresses the underlying sharing scenario while the second part addresses the combined sharing scenario (impact of BSU probabilistic activity). In the first part, we benchmark our work with respect to three schemes. That is, optimizing the problem

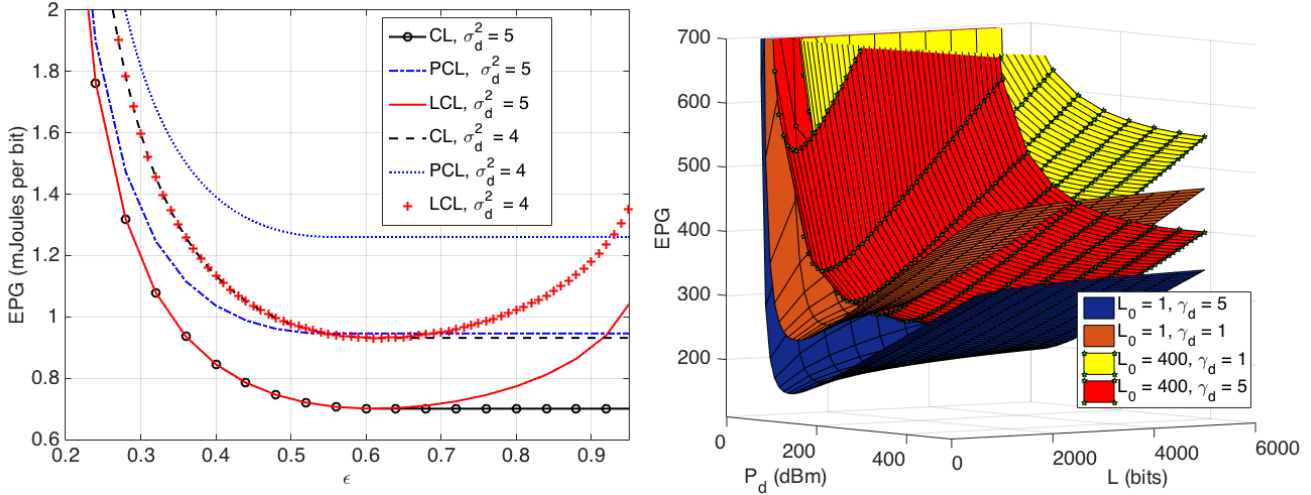
with respect to only the transmission power, denoted as PCL, and optimizing the problem with respect to only the frame length, denoted as LCL, subject to the previously mentioned constraint. In the last comparison scheme we utilize the same methodology developed in [10], denoted as NCL. This scheme does not consider the impact of retransmission power consumption or frame length in the problem design. We then show the improvement gain by jointly optimizing both variables, transmission power and frame length (noted as CL), in comparison to all benchmark schemes. In the second part, we evaluate the combined sharing scheme, denoted as CS, with respect to underlying sharing, denoted as US, and opportunistic sharing, denoted as OS. Note that we consider that all channel gains follow an exponential distribution.

TABLE II
SIMULATION PARAMETERS.

Parameter Name	Value
Wireless channels	Rayleigh Fading
P_b, P_{pk}	20 dBm
ϵ	0.6
Primary Rate (R_b)	0.5 symbol / sec
$\sigma_{db}^2, \sigma_{bd}^2$	1
σ_d^2	1, 4, 5

Figure 3(a) evaluates EPG versus ϵ under several optimization variables and different DDU's channel's parameters ($\sigma_d^2 = 4, \sigma_d^2 = 5$). Figure 3(a) considers the underlying scenario. It is clear that better channel quality (higher σ_d^2) results in an improved EPG. Increasing ϵ , hence increasing Q_I , results in improving EPG up to a certain value Q_I , where $P_{pk} < Q_I$. It is noted that CL achieves better performance than LCL and PCL for all values of ϵ . To explain the behavior of LCL performance with respect to ϵ , it is essential to note that in LCL we increase P_d with the increase of ϵ . This design helps reflecting the impact of ϵ on LCL performance. This design improves the LCL performance up to a certain optimal P_d (associate with a certain ϵ). Then the LCL performance starts degrading.

In underlying scenario, Fig. 3(b) evaluates the MEPG metric versus both P_d and L , under several MAC layer overhead, L_0 , and channel quality, γ_d . This figure verifies the strict quasi-convexity structures of the MEPG with respect to each variable. From Fig. 3(b) it is clear that MEPG has a unique joint global minimum for all P_d and L which aligns with the analytical



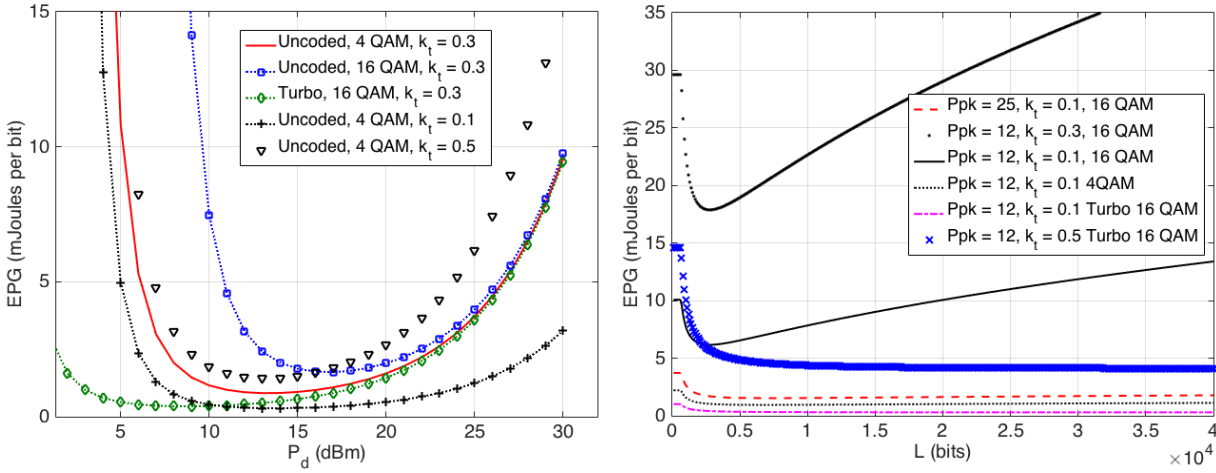
(a) EPG performance for different schemes versus ϵ for $\sigma_d^2 = 4$ and $\sigma_d^2 = 5$. (b) \mathcal{E}_e performance versus P_d and L given different channel quality, γ_d , and MAC overhead (L_0).

Fig. 3. EPG vs ϵ and a surface plot of EPG against the P_d and L .

result in the previous section. We also observe that changing L_0 has a different impact on the structure of EPG in compared to changing γ_d . It is noted that the structure of EPG for $L_0 = 400$ and $L_0 = 1$ are similar to that in Fig. 4(a) and Fig. 4(b).

Figure 4(a) evaluates the EPG performance of LCL scheme versus P_d , using optimal L , under several coding and modulation schemes, i.e., Uncoded, Turbo coded, 4 QAM and 16 QAM, and several values of $k_t = 0.1, 0.3, 0.5$. In Fig. 4(a), we consider underlying scenario. It is noted that the Turbo coded communication achieves lower EPG in comparison to the uncoded communication. We also observe that higher modulation constellation (16 QAM) achieves worse EPG compared to the lower modulation constellation (4 QAM). However, at high value of the transmission power, all schemes approach similar EPG performance. Increasing k_t increases the value of EPG. It is also noted that low values of k_t increases the tolerance of EPG toward the increment of P_d .

Figure 4(b) evaluates the EPG versus L , using optimal P_d , in underlying scenario, several coding, and modulation schemes, i.e., Uncoded, Turbo coded, 4 QAM and 16 QAM, and several values of k_t . The change of the EPG structure is observed via changing the modulation scheme and power constraint P_{pk} . Inline with the results in Fig. 4(a), it is noted that the Turbo coded communication achieves lower EPG in comparison to the uncoded communication. We also



(a) EPG performance of LCL scheme versus P_d and several (b) EPG performance of PCL versus the frame length, for several modulation, coding schemes, k_t , and P_{pk} .

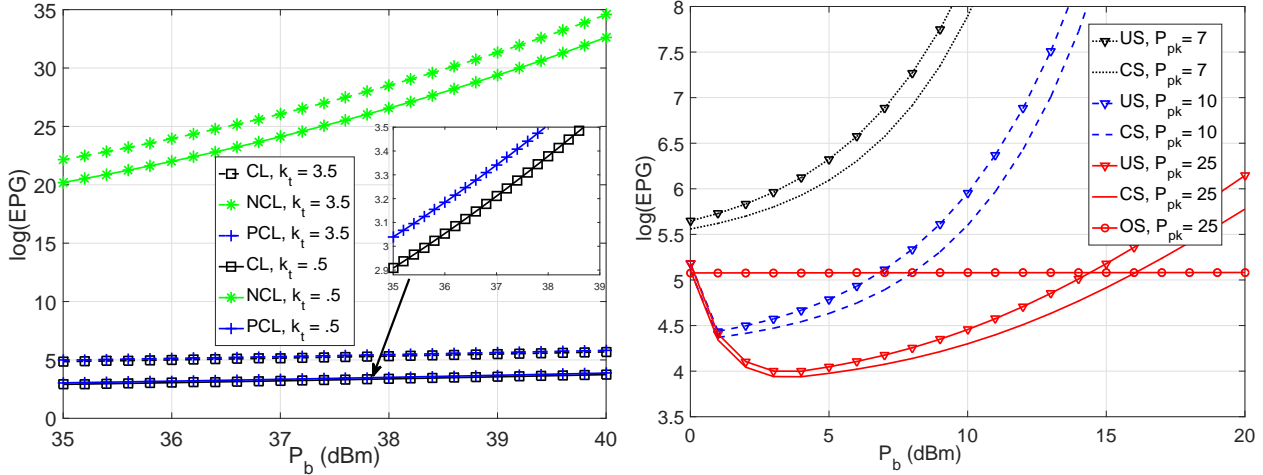
Fig. 4. EPG vs P_d for LCL scheme and EPG vs L for PCL scheme.

observe that higher modulation constellation (16 QAM) achieves worse EPG compared to the lower modulation constellation (4 QAM). Increasing k_t power increases the value of EPG.

We introduce Fig. 5(a) to compare between [10] and our proposed framework. In this figure, we evaluate the EPG performance for three different schemes versus BSU's transmission power with different circuit power (k_t). All schemes CL, PCL, and NCL have already been defined at the beginning of this section. Note that we present the logarithm of EPG (not only EPG) because the difference in variation of all schemes is so large to be presented in the same figure. It is observed that CL algorithm outperforms both PCL and NCL. Also, scheme PCL outperforms scheme NCL. This is because PCL consider the retransmission power in formulating the optimization framework, whereas, NCL does not consider the retransmission power.

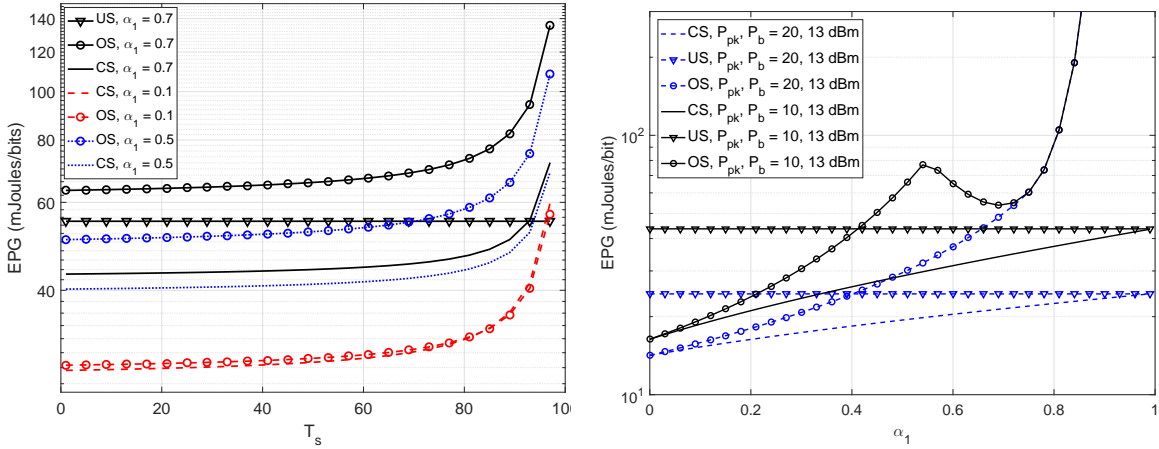
In the below figures, Fig. 5(b), Fig. 6(a), Fig. 6(b), Fig. 7(a), Fig. 7(b), and Fig. 8, we evaluate the combined sharing scheme (with sensing information about BSU activities) proposed in Sec. V versus two others sharing scheme, i.e. underlying sharing (in Sec. III) and opportunistic sharing. The combined sharing scheme is denoted as CS, the underlying sharing is denoted as US, whereas, the opportunistic sharing is denoted as OS.

In Fig. 5(b), we evaluate the EPG of all three schemes, i.e., CS, US, and OS, versus the BSU transmission power, P_b . As P_b represents an interference power, then it is expected that the EPG performance to be worse with the increase of P_b . Which is the observed behavior under $P_{pk} = 7$



(a) EPG comparison of our proposed framework versus [10], i.e., NCL, and P_b . (b) EPG performance of three sharing scenarios versus P_b for different P_{pk} .

Fig. 5. Comparison of EPG performance vs P_b for different schemes, CL, NCL, and PCL and different spectrum-sharing scenarios CS, US, and OS.



(a) EPG performance of three sharing scenarios versus T_s for different α_1 . (b) EPG performance of three sharing scenarios versus α_1 for different peak power values.

Fig. 6. Comparison of EPG performance vs T_s and α_1 for different spectrum-sharing scenarios, i.e., CS, US, and OS.

dBm and partial region when $P_{pk} = 10$ dBm and $P_{pk} = 25$ dBm. An important observation is the behavior of CS and US, when $P_b \leq 5$ dBm, for $P_{pk} = 10$ dBm and $P_{pk} = 25$ dBm. It is noted that EPG performance improves while the interference power is increased. This behavior is interpreted using the feasibility region concept of the optimization theory. That is, for small P_b the BSU outage probability is not met with huge DDU interference. Therefore, the interference constraint, in (18b), must be tight. Increasing P_b will relax the interference constraint, thus enlarge the feasibility region, therefore, EPG improves with the increase of P_b .

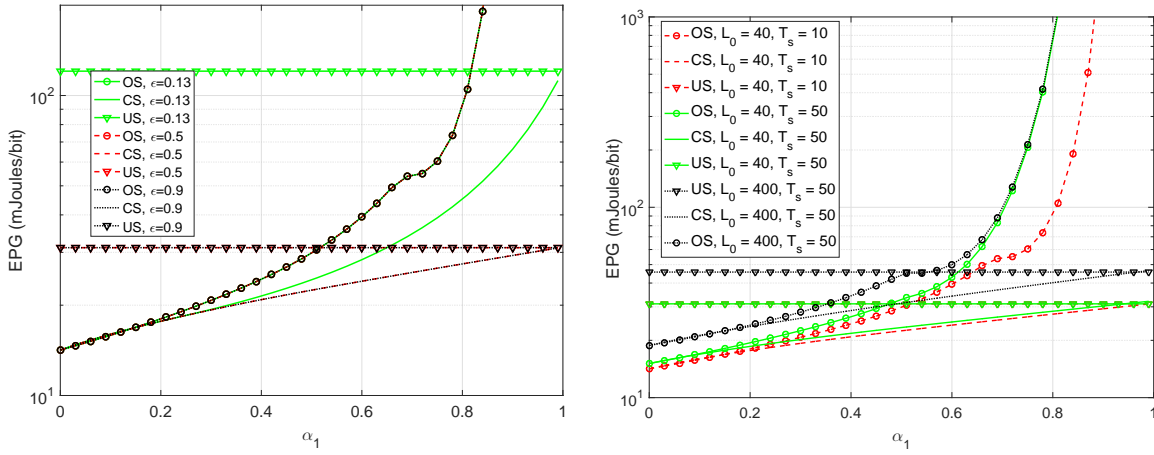
In Fig. 6(a), we evaluate the EPG performance of two sharing scenarios, CS and OS, versus

T_s for different values of $\alpha_1 = 0.1, 0.5, 0.7$. As expected from analytical results, the higher the sensing time the worse the EPG performance of both evaluated schemes, given that the impact of sensing time on sensing scheme efficiency is not considered. The main observation in Fig. 6(a) is about the difference in performance between CS and OS. It is noted that as $\alpha_1 \rightarrow 1$ CS outperforms OS with larger gap. Whereas, if $\alpha_1 \rightarrow 0$ the difference gap between OS and CS is almost zero.

Fig. 6(b) shows the EPG performance of DDU versus α_1 for several peak transmission power values, i.e., $P_{pk} = 10, 20$, $P_b = 13$ dBm. In general, the CS EPG performance lower bounds both OS and US EPG performances. It is noted that as $\alpha_1 \rightarrow 1$ the CS performance converges to the performance of US, where both of CS and US performances considerably outperforms that of OS. However, as $\alpha_1 \rightarrow 0$, BS is idle all the time, the performance of CS converges to that of OS, where both CS and OS outperforms US. It is noted that the US performance does not change with the change of α_1 because it assumes that BSU is always active. It is observed that there is a peak of OS's EPG performance around $\alpha_1 \approx 5.2$. This peak occurs because the QoS constraint, corresponding to (18c), which acts as a lower bound constraint of the OS system, conflicts with the peak power constraint at this value of α_1 . Therefore, to solve this in-feasibility problem, we ignored the peak power constraint, because in OS increasing the power does not harm the BSU. Hence, for high α_1 we note that the performance of OS with $P_{pk} = 10$ dBm converges to that of OS with $P_{pk} = 20$ dBm, where there is no conflict between the QoS and peak power constraint since $P_{pk} = 20$ is high enough to support the minimum QoS constraint.

Fig. 7(a) shows the EPG performance of DDU versus α_1 for several interference tolerance threshold values, i.e., $\epsilon = 0.13, 0.5, 0.9$. The performance of OS does not vary as the value ϵ changes. This is expected as OS scheme enables DDU to operate only when BSU is idle, thus does not consider ϵ into account. As expected from analytical results, the performance of CS scheme converges to that of US as $\alpha_1 \rightarrow 1$. Where the performance of CS scheme converges to that of OS as $\alpha_1 \rightarrow 0$. As ϵ decreases the interference constraint become stricter and thus the OS curve intersect with that of US at higher value of α_1 .

Finally, Fig. 7(b) evaluates the EPG performance of DDU versus α_1 for several sensing time



(a) EPG performance of three sharing scenarios versus α_1 for different tolerance threshold. (b) EPG performance of three sharing scenarios versus α_1 for different sensing time and overhead length.

Fig. 7. Comparison of EPG performance vs α_1 for different values of ϵ , T_s , and L_0 , under CS, US, and OS sharing scenarios. ($T_s/T_c = 0.1$ and 0.5 , given $T_c = 100$) and overhead length ($L_0 = 40, 400$). Similar to other simulation results, CS scheme lower bounds both US and OS for different values of T_s and L_0 . Also, CS converges to US as $\alpha_1 \rightarrow 1$ and it converges to OS as $\alpha_1 \rightarrow 0$ (BS is idle). An interesting observation is that increasing T_s considerably impacts EPG performance at high values of α_1 , whereas, increasing L_0 impact the EPG performance at lower values of α_1 .

The impact of guaranteeing minimum throughput on numerical results is described in Fig. 8, under three sharing scenarios, US, OS, and CS. The EPG versus minimum rate threshold for several BSU's outage tolerance threshold, $\epsilon = 0.6, 0.7, 0.9$ is evaluated. It is observed that for all three scenarios the left side of the curves are not changed upto a certain R_s . This is interpreted by the fact that for small values of R_s the minimum rate constraint is not active, hence the EPG performance is not changed. Increasing R_s more than a certain value results in increasing/degrading the EPG performance of all three scenarios, US, OS, and CS. This is an expected result as increasing R_s will reduce the feasibility region, hence, the performance will degrade. Finally, it is observed that the EPG performance of both US and CS saturates again at high values of R_s . The reason for this saturation is that the lower bound, which is produced by the minimum rate constraint, is equal or larger than the upper bound of the problem, which is produced by the peak power constraint P^{st} . In order to protect the BSU minimum outage constraint, which is merged into the peak power constraint, we enforce the peak power constraint as the active constraint in this conflict between the lower and upper bound constraints, which

results in inactive QoS constraint. Note that, since OS does not need to protect the BSU's minimum QoS, thus, the OS EPG performance is always increasing.

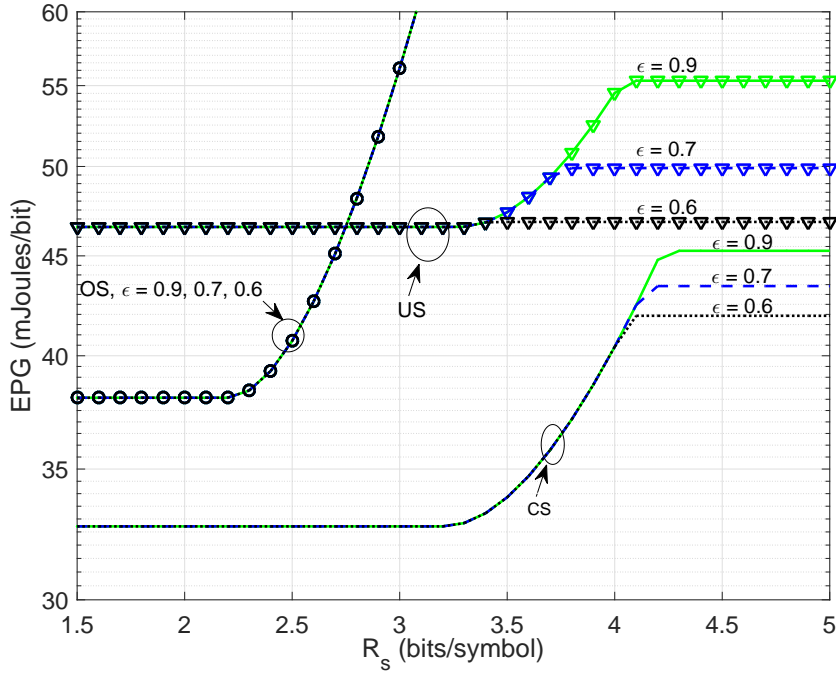


Fig. 8. EPG performance of three sharing scenarios versus minimum rate threshold.

VII. CONCLUSION

In this work, we considered a spectrum-sharing model where we minimized the energy per good bit of the device-to-device users while protecting the base station associated users. A strict pseudo-convex structure of the problem is verified. Analytical expressions of the optimal power and frame length were provided. It is shown that optimizing the problem with respect to the power and the frame length, considerably, improves the system's performance in compared to conventional methods, by about 0.4 Joules per bit. In later sections, we proposed a combined sharing scheme to improve the energy efficiency performance. Analytical and numerical results are provided to evaluate the proposed combined sharing scheme versus underlying and opportunistic sharing schemes.

APPENDIX A

PROOF OF MUTUAL INFORMATION W/O CSI.

Let us note the received signal at the DDU's receiver is expressed as,

$$Y = h_d X_d + h_{bd} X_b + N_d \quad (25)$$

where X_d and X_b are the secondary user (DDU) and primary user (BSU) signals, respectively, and N_s is the associated noise. Our target is to obtain the mutual information of the secondary user conditioning on the availability of the secondary channel h_d , while CSI of primary to secondary channel h_{bd} is not available. The mutual information of the secondary user is obtained as,

$$\mathcal{I}(Y; X_d | h_d) = h(X_d | h_d) - h(X_d | Y, h_d) \quad (26)$$

where $\mathcal{I}(Y; X_d | h_d)$ is the conditional mutual information between Y and X_d . The function $h(X_d | h_d)$ expresses the entropy of random variable X_d given h_d . We use the fact that adding a constant to a random variable does not change its entropy and the fact that conditioning always reduces the entropy. We also consider that the entropy of any random variable is always upper-bounded by the entropy of Gaussian random variable of the same variance. It follows that,

$$\begin{aligned} h(X_d | Y, h_d) &= h(X_d - \eta Y | Y, h_d) \leq h(X_d | Y = y, h_d) \\ &= h(X_d - \eta Y | h_d) \leq \frac{1}{2} \log(2\pi e \text{Var}(X_d - \eta Y)) \end{aligned} \quad (27)$$

where $\text{Var}(\cdot)$ is the variance of its argument. We now find η so that ηY is the minimum linear mean-square error estimate of X_d [32]. It follows that,

$$\eta = \frac{\mathbb{E}[X_d Y | h_d]}{\mathbb{E}[Y^2 | h_d]} = \frac{h_d P_d}{\gamma_d P_d + \sigma_{bd}^2 P_b + \sigma_n^2} \quad (28)$$

where σ_n^2 is the noise variance. It follows that the variance of $X_d - \eta Y$ is expressed as follows,

$$\text{Var}(X_d - \eta Y | h_d) = \frac{P_d P_b \sigma_{bd}^2 + P_d \sigma_n^2}{\gamma_d P_d + \sigma_{bd}^2 P_b + \sigma_n^2} \quad (29)$$

We then find the upper bound of the entropy as follows,

$$h(X_d|Y, h_d) \leq \log \left(2\pi e \frac{P_d P_b \sigma_{bd}^2 + P_d \sigma_n^2}{\gamma_d P_d + \sigma_{bd}^2 P_b + \sigma_n^2} \right) \quad (30)$$

Finally, the targeted mutual information is expressed as follows,

$$\begin{aligned} \mathcal{I}(X_d|Y, h_d) &\geq \log(2\pi e P_d) - \log \left(2\pi e \frac{P_d P_b \sigma_{bd}^2 + P_d \sigma_n^2}{\gamma_d P_d + \sigma_{bd}^2 P_b + \sigma_n^2} \right) \\ &\geq \log \left(1 + \frac{\gamma_d P_d}{P_b \sigma_{bd}^2 + \sigma_n^2} \right) \end{aligned} \quad (31)$$

The term in (31) is the one we used in the denominator of (1a).

APPENDIX B

EXISTENCE OF GLOBAL MINIMUM OF PROPOSITION 1.

It is clear that $\nabla \mathcal{E}_e(P_d^*) = 0$ iff $\nabla \mathcal{E}_n(P_d^*) \mathcal{E}_d(P_d^*) - \mathcal{E}_n(P_d^*) \nabla \mathcal{E}_d(P_d^*) = 0$. Then, $\nabla \mathcal{E}_n(P_d^*) = \mathcal{E}_e(P_d^*) \nabla \mathcal{E}_d(P_d^*)$. Utilizing the properties of \mathcal{E}_n function,

$$\mathcal{E}_n(P_d) = \mathcal{E}_n(P_d^*) + \nabla \mathcal{E}_n(P_d^*) (P_d - P_d^*) \quad (32a)$$

$$\geq \mathcal{E}_n(P_d^*) + \mathcal{E}(P_d^*) (\mathcal{E}_d(P_d) - \mathcal{E}_d(P_d^*)) \quad (32b)$$

$$= \mathcal{E}_n(P_d^*) + \mathcal{E}(P_d^*) \mathcal{E}_d(P_d) - \mathcal{E}_n(P_d^*) \implies \frac{\mathcal{E}_n(P_d)}{\mathcal{E}_d(P_d)} \geq \mathcal{E}(P_d^*), \quad (32c)$$

(32a) follows from the convex property of \mathcal{E}_n and (32b) is due to the concave structure of $\mathcal{E}_d(P_d)$.

APPENDIX C

PROOF OF PROPOSITION 2.

The strict quasi-convexity structure of $\mathcal{E}_e(L) = \mathcal{E}_n^{(1)} + \mathcal{E}_n^{(2)}$ is verified by showing that $\mathcal{E}_e(L)$ is a decreasing function or increasing function or decreasing then increasing function on L [28]. Let us note $\mathcal{E}_e \uparrow$ for an increasing \mathcal{E}_e and $\mathcal{E}_e \downarrow$ for a decreasing \mathcal{E}_e . The behavior of $\mathcal{E}_e(L)$ is studied through the following inequality,

$$\frac{\partial \mathcal{E}_e(L)}{\partial L} = \frac{k'_M}{P_d} L_F L - 2L_F L_0 + \frac{k'_M}{P_d} L^2 - L_0 L \begin{matrix} \mathcal{E}_e \uparrow \\ \mathcal{E}_e \downarrow \end{matrix} \gtrless 0. \quad (33)$$

Note that as $L \rightarrow \infty$ the function \mathcal{E}_e is increasing function,

$$\lim_{L \rightarrow \infty} \frac{\partial \mathcal{E}_e(L)}{\partial L} > 0. \quad (34)$$

However, as $L \rightarrow L_0$, then,

$$\begin{aligned} \frac{k'_M}{P_d} > \frac{5}{3} &\implies \mathcal{E}_e(L) \text{ is increasing} \\ \frac{k'_M}{P_d} < \frac{5}{3} &\implies \mathcal{E}_e(L) \text{ is decreasing} \end{aligned} \quad (35)$$

For general L , it is noted that,

$$\begin{aligned} L \in [0, L_2^*] &\implies \mathcal{E}_e(L) \text{ is decreasing} \\ L \in [L_2^*, \infty] &\implies \mathcal{E}_e(L) \text{ is increasing} \end{aligned} \quad (36)$$

where $L_2^* = L_0 \left[3 - \frac{k'_M}{P_d} + \sqrt{\left(\frac{k'_M}{P_d}\right)^2 + 10\frac{k'_M}{P_d} + 9} \right]$. Therefore, combining (34), (35), and (36) it is clear that $\mathcal{E}_e(L)$ is a strictly decreasing function then strictly increasing function. Thus, $\mathcal{E}_e(L)$ is a strictly quasi-convex with respect to L . Since $\frac{\partial \mathcal{E}_e(L)}{\partial L} = 0$ at L_2^* , then, $\mathcal{E}_e(L)$ is also a pseudo-convex [28]. Since $\mathcal{E}_e(L)$ is both strictly quasi-convex and pseudo-convex, then $\mathcal{E}_e(L)$ is strictly pseudo-convex.

APPENDIX D

TRANSFORMATION OF PROBLEM (18), LEMMA 2.

We begin the transformation proof (equivalence proof between problem (18) and problem (19)) by applying the fractional/parametric optimization technique on the objective function (18a) [30]. Problem (18) is then reformulated to the following problem,

$$\min_{P_d, L, t} \quad g_{pa}(P_d, L, t) \quad (37a)$$

$$\text{s.t.} \quad P_d \leq \min \{Q_I, P_{pk}\} \quad (37b)$$

$$\frac{T_c}{T} \left[\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d) \right] \geq R_s, \quad (37c)$$

where $g_{pa}(P_d, L, t) = \frac{T_c L_F}{T L} \left\{ \alpha_1 e^{\left(\frac{a'(L)}{P_d}\right)} + \alpha_0 e^{\left(\frac{a(L)}{P_d}\right)} \right\} \left(\frac{L_F}{L} k_t P_d + k_c \right) + P_{sns} \frac{T_s}{T} - t \frac{T_c}{T} \left(\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log(1 + P_d \gamma_d) \right)$.

Because of the complexity of (37a), as a function of P_d , the generalized convexity theory is not insightful in obtaining the global optimal solution. Therefore, we apply a change of variable technique and introduce a new equality constraint to compensate for the added new variable. It

follows that the optimization problem in (37) is reformulated as follows,

$$\min_{P_d, P_{de}, L, t} \left[\alpha_1 e^{\left(\frac{a'(L)}{P_{de}}\right)} + \alpha_0 e^{\left(\frac{a(L)}{P_{de}}\right)} \right] \frac{T_c L_F}{T L} \left[\frac{L_F}{L} k_t P_d + k_c \right] + P_{sns} \frac{T_s}{T} - \frac{t T_c}{T} \left[\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d) \right] \quad (38a)$$

$$\text{s.t.} \quad P_d \leq \min \{Q_I, P_{pk}\} \quad (38b)$$

$$\frac{T_c}{T} \left[\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d) \right] \geq R_s \quad (38c)$$

$$P_d = P_{de} , \quad (38d)$$

where P_{de} is the new optimization variable. Constraint (38d) is included to guarantee the equivalence between both P_d and P_{de} . The above problem is exactly equivalent to problem (19). This decomposition reduces the complexity of the problem as a function of P_d .

APPENDIX E

STRUCTURE VERIFICATION OF PROBLEM (19), LEMMA 3.

To verify the strict quasi-convexity structure of problem (19), it is enough to verify that the structure of its objective function, i.e., (19a), is strictly quasi-convex with respect to each and all variables, i.e., $\{t, P_d, P_{de}, L\}$. We begin by the verifying the strict quasi-convexity of (19a) with respect to t . This is straight forward, since (19a) is a linear function of t . Then, the strict quasi-convexity of the (19a) with respect to P_d can be inferred following similar lines as in Proposition 1, while noting that the weighted sum of strict concave functions is still strict concave. The structure of (19a) with respect to P_{de} is found by noting that the only dependent term on P_{de} is $\left[\alpha_1 e^{\left(\frac{a'(L)}{P_{de}}\right)} + \alpha_0 e^{\left(\frac{a(L)}{P_{de}}\right)} \right]$. We also observe that this term is monotonically decreasing with respect to P_{de} . Thus, (19a) is strict quasi-convex with respect to P_{de} .

We now check the objective function structure with respect to L . We note that only the first term of (19a) depends on L . Thus, it is sufficient to show that the first term is strict quasi-convex with respect to L . We first apply transformation of the first term of the objective function by using logarithm function, since the transformation function α is a monotonic increasing function then the both problems before and after transformation are equivalent [33]. The transformed

function is expressed as follows,

$$g'_{pa}(L) = \log(\alpha_0 c) + \beta \log(L_F) + \log \left[\frac{\alpha_1 c'}{\alpha_0 c} e^{\beta'' \log L_F} + 1 \right] + 2 \log(L_F) - 2 \log(L) + \log(T_c) - \log(T), \quad (39)$$

where $c = e^{\frac{b_M}{P_d}} k_t P_d$, $c' = e^{\frac{b'_M}{P_d}} k_t P_d$, $\beta = \frac{k_M}{P_d}$, and $\beta'' = \frac{k'_M}{P_d} - \frac{k_M}{P_d}$ are positive constants that do not depend on L . The designated $g'_{pa}(L)$ function is the transformed first term of $g_{pa}(L)$, via a logarithm function. Note that in the formulation of $g'_{pa}(L)$ we assume that $\frac{L_F}{L} k_t P_d \gg k_c$. The partial first derivative of $g'_{pa}(L)$ with respect to L is expressed as follows,

$$\frac{\partial g'_{pa}(L)}{\partial L} = \frac{\beta}{L_F} + \frac{\frac{\alpha_1 c' \beta''}{\alpha_0 c L_F} e^{\beta'' \log L_F}}{\frac{\alpha_1 c'}{\alpha_0 c} e^{\beta'' \log L_F} + 1} + \frac{2}{L_F} - \frac{2}{L}, \quad (40)$$

In similar lines with Proposition 2 we show the structure of (19a) with respect to L by studying the behavior of the function at $L \rightarrow L_0$ and $L \rightarrow \infty$ as follows,

$$\text{At : } L \rightarrow L_0 \quad \begin{cases} \text{If } \frac{\beta}{2L_0} + \frac{\frac{\alpha_1 c' \beta''}{\alpha_0 c L_F} e^{\beta'' \log L_F}}{\frac{\alpha_1 c'}{\alpha_0 c} e^{\beta'' \log L_F} + 1} < \frac{1}{L_0} \implies g_{pa}(L) \downarrow \\ \text{If } \frac{\beta}{2L_0} + \frac{\frac{\alpha_1 c' \beta''}{\alpha_0 c L_F} e^{\beta'' \log L_F}}{\frac{\alpha_1 c'}{\alpha_0 c} e^{\beta'' \log L_F} + 1} > \frac{1}{L_p} \implies g_{pa}(L) \uparrow \end{cases}. \quad (41)$$

On the other hand, we observe that as $L \rightarrow \infty \implies g_{pa} \uparrow$. Therefore, we note that $g_{pa}(L)$ is a decreasing then increasing function or always increasing function on L . This conclude the proof of the quasi-convexity of $g_{pa}(L)$ with respect to L .

APPENDIX F

PROOF OF PROPOSITION 3.

In this appendix, we find the optimal frame length that minimize problem (18), as in (20).

Utilizing the fractional programming we change the structure of the objective function as,

$$\begin{aligned} g'_{pa}(L) &= \left[\alpha_1 e^{\left(\frac{a'(L)}{P_d}\right)} + \alpha_0 e^{\left(\frac{a(L)}{P_d}\right)} \right] L_F^2 k_t P_d - t L^2 \Gamma \\ &= \left[\alpha_1 c' (L_F)^{\beta'+2} + \alpha_0 c (L_F)^{\beta+2} \right] - t L^2 \Gamma \end{aligned} \quad (42)$$

where $c' = e^{\frac{b'_M}{P_d}} k_t P_d$ and $c = e^{\frac{b_M}{P_d}} k_t P_d$.

The optimization of problem (43) is not straight forward, because of the different geometric and exponential structures of the objective as a function of the frame length. To overcome this

problem, we apply several procedures. We begin by introducing new optimization variables with the associated equality constraints and using the reformulated objective function in (42) as,

$$\min_{L_F, L_p, L} \quad \alpha_1 c' (L_p)^{\beta'+2} + \alpha_0 c (L_F)^{\beta+2} - tL^2\Gamma \quad (43a)$$

$$\text{s.t.} \quad L_F = L_p \quad (43b)$$

$$L + L_0 = L_F \quad (43c)$$

We then minimize problem (43) with respect to each variable $\{L_F, L_p, L\}$. It is observed that the objective function (43a) is a linear combination of three functions, i.e., $g_a(L_p) = \alpha_1 c' (L_p)^{\beta'+2}$, $g_b(L_F) = \alpha_0 c (L_F)^{\beta+2}$, and $g_c(L) = -tL^2\Gamma$. It is also noted that each of these function depends on one of the optimization variables, i.e., $g_a(L_p)$ depends only on L_p , $g_b(L_F)$ depends only on L_F , and $g_c(L)$ depends only on L . This argument leads to the fact that minimizing (43a) with respect to L_p is equivalent to minimizing $g_a(L_p)$ with respect to L_p , minimizing (43a) with respect to L_F is equivalent to minimizing $g_b(L_F)$ with respect to L_F , and finally minimizing (43a) is equivalent to minimizing $g_c(L)$ with respect to L . All the above mentioned three optimization problem preserve both equality constraints, i.e., (43b) and (43c). Let us designate the above three sub-problems as:

- $\mathfrak{P}_a : \min_{L_p} \quad g_a(L_p) \text{ s.t.: (43b) \& (43c),}$
- $\mathfrak{P}_b : \min_{L_F} \quad g_b(L_F) \text{ s.t.: (43b) \& (43c),}$
- $\mathfrak{P}_c : \min_L \quad g_c(L) \text{ s.t.: (43b) \& (43c).}$

Both problems \mathfrak{P}_a and \mathfrak{P}_b consist of geometrical expressions in addition to the two equality constraint. Thus, to solve the above problems, we use a transformation of objective theory, i.e., $\min_x g(x)$ is equivalent to $\min_x \phi(g(x))$ given that $\phi(\cdot)$ is an $\mathbb{R} \rightarrow \mathbb{R}$ monotonic increasing function [33]. Hence, we use a logarithm function (as $\phi(\cdot)$) on the objective functions, i.e., $\log g_a(L_p)$ and $\log g_b(L_F)$. It follows that the optimal solutions of the three sub-problems, i.e., \mathfrak{P}_a , \mathfrak{P}_b , and \mathfrak{P}_c are expressed, respectively, as follows, $\hat{L}_F = \frac{\beta+2}{\mu_1+\mu_2}$, $\hat{L}_p = \frac{\beta'+2}{\mu_1}$, and $L = \frac{\mu_2}{2t\Gamma}$, where μ_1 and μ_2 are the associated Lagrangian multipliers of both (43b) and (43c), respectively. Solving for the optimal value of both Lagrangian multipliers we obtained, $\hat{\mu}_2 = \hat{\mu}_1 \frac{\beta-\beta'}{\beta'+2}$ and $\hat{\mu}_1 = \frac{\beta'+2}{\beta-\beta'} \left[-L_0 t\Gamma + \sqrt{L_0^2 t^2 \Gamma^2 + 2(\beta-\beta') t\Gamma} \right]$. We then substitute $\hat{\mu}_1$ and $\hat{\mu}_2$ in \hat{L}_F , \hat{L}_p , and

L . It is found that optimal values of all variables are equal, i.e., $L_F = L_p = L + L_0$. It follows that the optimal frame length that minimizes (19) is expressed as,

$$\hat{L}_F = \frac{\beta' - \beta}{\left[L_0 t \Gamma - \sqrt{L_0^2 t^2 \Gamma^2 + 2(\beta - \beta') t \Gamma} \right]} \quad (44)$$

which is similar to that of \hat{L}_F in Proposition 3.

APPENDIX G

PROOF OF PROPOSITION 4.

We begin the proof of Proposition 4 by constructing the Lagrangian function of problem (19) as a function of both P_d and P_{de} .

$$\begin{aligned} \mathfrak{L} = & \left[\alpha_1 e^{\left(\frac{\alpha'(L)}{P_{de}}\right)} + \alpha_0 e^{\left(\frac{\alpha(L)}{P_{de}}\right)} \right] \frac{T_c L_F}{LT} \left[\frac{L_F}{L} k_t P_d + k_c \right] \\ & - \frac{t T_c}{T} \left[\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d) \right] + P_{sns} \frac{T_s}{T} \\ & + \lambda \left[\frac{T_c}{T} \left(\alpha_1 \log \left(1 + \frac{P_d \gamma_d}{P_I} \right) + \alpha_0 \log (1 + P_d \gamma_d) - R_s \right) \right] + \mu [P_{de} - P_d] \end{aligned} \quad (45)$$

The optimal P_d that minimizes \mathcal{L} is found via finding the zero of the following function,

$$\begin{aligned} AP_d + \frac{T_c}{T} (\lambda - t) \left[\alpha_1 \frac{\gamma_d}{1 + P_I + P_d \gamma_d} + \alpha_0 \frac{\gamma_d}{1 + P_d \gamma_d} \right] - \mu = 0 & \iff \\ P_d^2 \frac{A}{B} \gamma_d^2 + \gamma_d P_d \left[\frac{A}{B} (1 + P_I) - 1 \right] + P_I \left[\frac{A}{B} - \alpha_0 \right] \alpha_1 = 0 \end{aligned} \quad (46)$$

where A and B are mentioned in Proposition 4. The result in (21) is found by algebraic manipulation of (46).

In order to find the optimal P_{de} which minimizes (19) we utilize several optimization techniques. We first observe that the first term of the objective function (19a) is the only term which depends on P_{de} . We also observe that the only dependent constraint on P_{de} is constraint (19d). Minimizing (19) with respect to only P_{de} , is equivalent to the following minimization problem,

$$\min_{P_{de}} \left[\alpha_1 e^{\left(\frac{\alpha'(L)}{P_{de}}\right)} + \alpha_0 e^{\left(\frac{\alpha(L)}{P_{de}}\right)} \right] J(P_d) \quad (47a)$$

$$\text{s.t.} \quad P_{de} = P_d \quad (47b)$$

where $J(P_d) = \frac{T_c L_F}{TL} \left[\frac{L_F}{L} k_t P_d + k_c \right]$. Note that the objective function of problem (47), without considering the equality constraint, is minimized by increasing P_{de} while maintaining a feasible

solution set. However, constraint (47b) relates the value of P_{de} and P_d through an equality, i.e., $P_{de} = P_d$. To solve for the optimal P_{de} , we divide problem (47) into two sub-problems and minimize each one of them with respect to an individual variable while introducing another equality constraint that equate between both new variables, as follows,

- $\min_{P_{d0}} \alpha_0 e^{\left(\frac{a(L)}{P_{d0}}\right)} J(P_d)$ s.t. $P_{d0} = P_d$ & $P_{d0} = P_{d1}$.
- $\min_{P_{d1}} \alpha_1 e^{\left(\frac{a'(L)}{P_{d1}}\right)} J(P_d)$ s.t. $P_{d0} = P_d$ & $P_{d0} = P_{d1}$.

We then apply a transformation of objective theory, i.e., $\min_x g(x)$ is equivalent to $\min_x \phi(\gamma(x))$ given that ϕ is an $\mathbb{R} \rightarrow \mathbb{R}$ monotonic increasing function [33]. We select $\log(\cdot)$ as the transformation function. Solving both sub-problems the optimal P_{de} is expressed as,

$$\hat{P}_{de} = \sqrt{\frac{a(L) + a'(L)}{\mu}}, \quad (48)$$

which is equivalent to that in (22). This expression is interpreted as the relation between \hat{P}_{de} to \hat{P}_d through μ has huge impact on the value of optimal P_{de} . This is an intuitive solution since the objective of (47) is a decreasing function of P_{de} . Thus, given a relaxed equality constraint (small μ) we can increase P_{de} as high as possible, while respecting P_d .

REFERENCES

- [1] 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.
- [2] 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.
- [3] X. Ge, S. Tu, G. Mao, C. X. Wang, and T. Han, "5G ultra-dense cellular networks," *IEEE Wireless Communications*, vol. 23, no. 1, pp. 72–79, February 2016.
- [4] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 30–37, June 2011.
- [5] S. Zhang, Q. Wu, S. Xu, and G. Li, "Fundamental green tradeoffs: Progresses, challenges, and impacts on 5G networks," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–1, 2016.
- [6] Z. Zhou, M. Dong, K. Ota, J. Wu, and T. Sato, "Energy efficiency and spectral efficiency tradeoff in device-to-device (D2D) communications," *CoRR*, vol. abs/1407.1556, 2014. [Online]. Available: <http://arxiv.org/abs/1407.1556>
- [7] E. Yaacoub and O. Kubbar, "Energy-efficient device-to-device communications in LTE public safety networks," in *2012 IEEE Globecom Workshops*, Dec 2012, pp. 391–395.

- [8] A. Alabbasi and B. Shihada, "An energy efficient cognitive radio system with quantized soft sensing and duration analysis," in *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2015, pp. 423–428.
- [9] 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.
- [10] A. Alabbasi, Z. Rezki, and B. Shihada, "Energy efficient resource allocation for cognitive radios: A generalized sensing analysis," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2455–2469, May 2015.
- [11] C. Xiong, G. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 11, pp. 3874–3886, November 2011.
- [12] 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.
- [13] R. Zi, X. Ge, J. Thompson, C. X. Wang, H. Wang, and T. Han, "Energy efficiency optimization of 5G radio frequency chain systems," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 758–771, April 2016.
- [14] G. Li, Z. Hu, G. Zhang, L. Zhao, W. Li, and H. Tian, "Cross-layer design for energy efficiency of tcp traffic in cognitive radio networks," in *IEEE Vehicular Technology Conference (VTC Fall), 2011*, Sept 2011, pp. 1–5.
- [15] S. Althunibat, V. Sucasas, H. Marques, J. Rodriguez, R. Tafazolli, and F. Granelli, "On the trade-off between security and energy efficiency in cooperative spectrum sensing for cognitive radio," *IEEE Communications Letters*, vol. 17, no. 8, pp. 1564–1567, August 2013.
- [16] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 12, no. 3, pp. 493–506, June 2004.
- [17] W. Su, S. Lee, D. Pados, and J. Matyjas, "Optimal power assignment for minimizing the average total transmission power in hybrid-arq Rayleigh fading links," *IEEE Transactions on Communications*, vol. 59, no. 7, pp. 1867–1877, July 2011.
- [18] J. Wu, G. Wang, and Y. Zheng, "Energy efficiency and spectral efficiency tradeoff in Type-I ARQ systems," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 356–366, February 2014.
- [19] J. Ren, Y. Zhang, N. Zhang, D. Zhang, and X. S. Shen, "Dynamic channel access to improve energy efficiency in cognitive radio sensor networks," *IEEE Transactions on Wireless Communications*, vol. PP, no. 99, pp. 1–1, 2016.
- [20] S. Bayhan and F. Alagoz, "Scheduling in centralized cognitive radio networks for energy efficiency," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 2, pp. 582–595, Feb 2013.
- [21] F. Haider, C. X. Wang, H. Haas, E. Hepsaydir, X. Ge, and D. Yuan, "Spectral and energy efficiency analysis for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 2969–2980, June 2015.
- [22] P. Serrano, A. Garcia-Saavedra, G. Bianchi, A. Banchs, and A. Azcorra, "Per-frame energy consumption in 802.11 devices and its implication on modeling and design," *IEEE/ACM Transactions on Networking*, vol. 23, no. 4, pp. 1243–1256, Aug 2015.
- [23] Y. Lin and V. Wong, "Wsn01-1: Frame aggregation and optimal frame size adaptation for ieee 802.11n w lans," in *IEEE Global Telecommunications Conference 2006, GLOBECOM '06*, Nov 2006, pp. 1–6.
- [24] A. Alabbasi and B. Shihada, "Energy efficient cross layer design for spectrum sharing systems," in *IEEE Wireless Communications and Networking Conference (WCNC), 2016*, April 2016.

- [25] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys Tutorials*, vol. 16, no. 4, pp. 1801–1819, Fourthquarter 2014.
- [26] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 933–946, 2000.
- [27] I. Chatzigeorgiou, I. Wassell, and R. Carrasco, "On the frame error rate of transmission schemes on quasi-static fading channels," in *42nd Annual Conference on Information Sciences and Systems, 2008. CISS*, March 2008, pp. 577–581.
- [28] A. Cambini and L. Martein, *Generalized Convexity and Optimization: Theory and Applications*. Springer, 2008.
- [29] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss-Seidel method under convex constraints," *Operations Research Letters*, vol. 26, no. 3, pp. 127 – 136, 2000.
- [30] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, 1967.
- [31] A. Daghistani, A. B. Khalifa, A. Showail, and B. Shihada, "Green partial packet recovery in wireless sensor networks," *Journal of Network and Computer Applications*, vol. 58, pp. 267 – 279, 2015.
- [32] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, New Jersey: Prentice-Hall, 1979.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*, 2004.