Mobile Networks and Applications (2021) 26:1750–1764
https://doi.org/10.1007/s11036-021-01755-z

Energy efficiency of full-duplex cognitive radio in low-power regimes under imperfect spectrum sensing

M. Ranjbar[1] · H. L. Nguyen[2] · N. H. Tran[1] · T. Karacolak[3] · S. Sastry[1] · L. D. Nguyen[4]

Accepted: 31 March 2021 / Published online: 28 April 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
This paper investigates the energy efficiency of a full duplex (FD) cognitive radio (CR) system in the low-power regime under a practical self-interference cancellation performance and imperfect spectrum sensing. We consider an opportunistic spectrum access network in which the secondary user (SU) is capable of self-interference suppression (SIS). The SIS ability enables the SU to work in simultaneous transmit and sense (TS) mode in order to increase the quality of channel sensing. Towards this goal, we first study the sensing performance, i.e., false-alarm and miss-detection probabilities, of the FD CR networks using TS, and compare the results with traditional half duplex (HD) CR systems using transmit only mode (TO). In the next step, we show that due to imperfect spectrum sensing, the secondary network channel is best characterized as a Gaussian-Mixture (GM) channel, which is widely used to capture the asynchronism in heterogeneous cellular networks. We then analytically characterize the low-signal-to-noise-ratio (low-SNR) metrics of the minimum energy per bit and wideband slope of the spectral-efficiency curve, and obtain these fundamental limits in closed-form. Furthermore, the characterization of these fundamental metrics allows us to identify practical signaling strategies that are optimally efficient in the low-SNR regime for the considered FD CR system. The benefits in terms of energy efficiency offered by FD CR over HD CR are also clearly demonstrated.

Keywords Energy efficiency · Full duplex radio · Cognitive radio · Spectrum sensing · Fundamental limits

1 Introduction

Recent years have witnessed an explosive growth in mobile data traffic generated by wireless devices, which is predicted to reach almost 50 billion users within the next decade. To keep pace with such demands, future generations of cellular networks are being designed to have multi-tier heterogeneous architectures involving macrocells, microcells, femtocells, device-to-device (D2D) links, Wi-Fi access points, and cognitive radio (CR) networks. Among these, CR has been considered as a revolutionary wireless communication paradigm due to its flexibility to improve spectrum efficiency by allowing cognitive users to access spectrum resources that are unoccupied by primary users (PUs) [1–7]. While research on CR is still an active area, the robust spectrum access offered by CR is undoubted, and it will directly benefit a broad spectrum of applications.

Among different access mechanisms, CR under the context of opportunistic spectrum access (OSA), i.e., interweave paradigm, has received significant attention (see [8–12] and references therein. In CR with OSA, the SU is compelled to monitor the availability of temporal white spaces before transmitting. This operation is referred as spectrum sensing. The literature on the analysis of OSA for improving spectrum efficiency and throughput/capacity in CR is vast, but mostly under the constraint of half-duplex (HD) communication, where a SU senses the channel before it can transmit. Given the recent development of several encouraging full-duplex (FD) designs to overcome the self-interference problem [13–16], the integration of full-duplex (FD) radio in CR has also been explored [17–30]. More specifically, the ability of FD cognitive users in sensing primary activity or receiving information while they are transmitting allows more flexibility in utilizing the spectrum holes and/or improving the throughput and reliability. For example, the ability of simultaneous transmit-and-sense

✉ H. L. Nguyen
nlhung@ac.udn.vn

Extended author information available on the last page of the article.

(TS) and simultaneous transmit-and-receive (TR) in FD CR has been investigated in [24, 25]. Under a similar TS framework, references [26, 27] proposed a "listen-and-talk" protocol for better spectrum utilization and sensing accuracy. In recent works in [28–30], CR networks assisted with FD relaying have also been investigated.

While spectrum sensing in CR is important, it is a challenging task and, unfortunately, it is never perfect in practice. Such imperfections must be taken taken into account in evaluating the sensing/throughput trade-off of the SU link. To date, existing works in the literature for both HD CR and FD CR with imperfect spectrum sensing rely on a traditional approach of treating aggregate interference from collisions due to the simultaneous spectrum access under occasional but consistent miss-detection as Gaussian [10, 11, 31]. While the Gaussian assumption greatly simplifies the analysis and has been prominent, this Gaussian assumption is unrealistic. For example, miss-detection occurs when the SU fails to detect an active PU at a given moment. In this event, PU transmissions generate interference to SU communication. The SU channel is therefore no longer Gaussian. In a similar manner, it has been shown in [32] that the use of finite-constellation inputs in primary networks leads to non-Gaussian interference plus noise that follows Gaussian mixture (GM) distributions in CR networks. In addition, it has been widely recognized that co-channel interference plus noise can be more accurately modeled as GM noise in heterogeneous wireless networks [33–36]. The realistic SU channel is therefore no longer Gaussian and the well-known expression log(1 + SNR) cannot be used. The consideration of realistic non-Gaussian co-channel interference plus noise will lead to a completely new link and network behavior of CR network, but it also poses difficult challenges. It is because even the information-theoretic limits for a basic communication link with non-Gaussian interference are not known [37–41]. Recently, the work in [42, 43] demonstrates that the optimal signaling scheme for a channel subject to GM noise is discrete in amplitude. However, numerical methods are still needed in order to compute the link capacity and the corresponding optimal input.

In wireless networks, especially CR, energy is at a premium, since energy resources are scarce and have to be conserved. As such, minimizing the energy cost per unit transmitted information will improve the efficiency. In the literature, the energy required to reliably send one bit is a widely adopted metric to measure the performance. Therefore, the energy efficiency is related to the fundamental limits through the channel capacity and optimal signaling schemes in the low-power regime. It is also worth mentioning that this low-power regime is the region in which both spectral efficiency and energy-per-bit are relatively low. As such, the low-power regime can also be referred to as the wideband regime [44–46]. There is a rich literature addressing fundamental limits on energy efficiency in CR, e.g., in [47–51], but mostly for half-duplex systems under the assumption of Gaussian-distributed aggregate interference. To our knowledge, no previous work in CR has examined energy efficiency in GM environments.

Motivated by the above observations, this paper attempts to develop new low-signal-to-noise-ratio (low-SNR) fundamental metrics of minimum energy per bit and wideband slope of the spectral-efficiency curve for CR systems under GM aggregate interference. The focus is on FD CR using simultaneous transmit-and-sense (TS) mode within the context of OSA, i.e., interweave paradigm, under the effect of miss-detection and false-alarm events due to imperfect spectrum sensing. For completeness of the solutions, we shall also compare with the traditional listen-before-talk operation in HD CR, which is referred to as transmit only (TO) mode. The developed information-theoretic findings on energy efficiency shall certainly provide useful results that lead to more efficient use of resources and shed light on the trade-off between bandwidth and power in CR networks.

In the first part of the paper, we first study the sensing performance, i.e., false-alarm and miss-detection probabilities, of the considered FD CR networks. The results are useful for the characterization of the low-SNR fundamental metrics of minimum energy per bit and wideband slope of the spectral-efficiency for the secondary network Gaussian-mixture channel under imperfect spectrum sensing. Instead of trying to maximize the mutual information or determine the optimal input directly, we will make use of the characterization that the first derivative of the capacity with respect to SNR at SNR$=0$ (or equivalently capacity per unit cost or minimum energy per bit when capacity is a concave function of SNR) can be determined from the Kullback-Leibler divergence without identifying the optimal input [52], [45, Equation (42)]. As a result, the minimum energy per bit and wideband slope can be obtained in closed-form. The characterization of these two fundamental metrics allows us to identify practical signaling strategies that are optimally efficient in the low-SNR regime for the considered FD CR system. The benefits in terms of energy efficiency offered by FD CR over HD CR are also clearly demonstrated.

2 System model

We consider a CR network in which a secondary user (SU) use the primary channels opportunistically; i.e SUs only start transmitting if the channel is sensed free. We assume that only one SU link is active at a time; i.e., SUs transmit using appropriate multiple access methods [53, 54].

In addition, the primary user (PU) activity is considered as an alternating ON/OFF random process with P_1 and P_0 as probability of primary user being ON and OFF, respectively [24–27].

2.1 Secondary channel model under imperfect FD spectrum sensing

The ability of sensing and transmitting simultaneously in FD allows more flexibility in detecting the presence of PUs. Specifically, as illustrated in Fig. 1, with FD, the SU transmits to its peer SU a data frame of length T, and, at the same time, it senses the channel for a potential transmission in the next frame [26, 27]. If the channel is sensed as busy at sensing decision time, the SU stops transmitting and only senses the channel until the next sensing decision time. The SU then starts transmitting again once the channel is sensed empty. This operation is referred to as simultaneous transmit-and-sense (TS) mode. In the case of traditional HD operation, sensing is initially performed in the first T_s seconds of a frame. If the channel is empty, the remaining time slot shall be used exclusively for data transmission, as shown in Fig. 1. If the channel is occupied, the SU is refrained from transmitting until the next sensing decision.

Depending on the sensing techniques applied, e.g, energy detection, a practical spectrum sensing performance can be characterized by the probabilities of miss-detection P_m and false alarm P_f. Because of the FD operation, these two parameters also depend on the quality of self-interference suppression (SIS), and will be analyzed in the next section. Now, let H_0 be the hypothesis that the channel is empty, and $\hat{H}_0$ be the hypothesis that SU transmitter perceives the channel to be free. On the other hand, H_1 is the hypothesis that the channel is busy, while $\hat{H}_1$ corresponds to the case that the channel is sensed as busy. Clearly, in the events of false alarm with the average probability P_f or correct detection with the average probability $1 - P_m$, the SU channel outputs are respectively noise and noise plus interference only (i.e., without SU transmission). Therefore, in those two events, channel erasure can be declared. Otherwise, SU transmits signal $\boldsymbol{x}$, and the output $\boldsymbol{y}$ is either

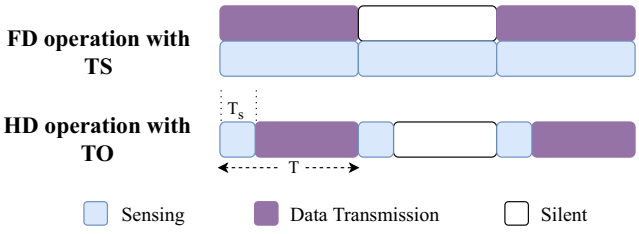

Fig. 1 FD operation with simultaneous transmit-and-sense (TS) mode versus HD operation with transmit only (TO) mode

$\boldsymbol{y} = \boldsymbol{h_s x} + \boldsymbol{n}$ under H_0 and $\hat{H}_0$, or $\boldsymbol{y} = \boldsymbol{h_s x} + \boldsymbol{n} + \boldsymbol{w}$ under H_1 and $\hat{H}_0$. Here, $\boldsymbol{n}$ is the Gaussian thermal noise with variance per real dimension $\frac{N_0}{2}$, $\boldsymbol{w}$ is PU interference, and $\boldsymbol{h_s}$ is the SU channel gain. Taking these events into consideration, the SU channel output can be written as

$$\boldsymbol{y} = b_e \boldsymbol{h_s x} + \boldsymbol{n} + \alpha \boldsymbol{w}. \tag{1}$$

In Eq. 1, b_e and α are binary indicator random variables with state space $\{0, 1\}$. $b_e = 0$ represents the event that SU transmitter perceives the time-slot to be busy, which includes both false alarm and correct-detection of an active PU, and $\alpha = 1$ denotes the event that PU is active; i.e., $P(\alpha = 1) = P_1$. In addition, $\alpha = 1$ given $b_e = 1$ represents the event that SU fails to detect an active PU, and the corresponding probability $P(\alpha = 1 \mid b_e = 1)$ can be calculated as:

$$\begin{aligned} P(\alpha = 1 \mid b_e = 1) &= 1 - P(\alpha = 0 \mid b_e = 1) = 1 - P\left(H_0 \mid \hat{H}_0\right) \\ &= 1 - \frac{\left(1 - P_f\right) P_0}{P_0\left(1 - P_f\right) + P_1 P_m}. \end{aligned} \tag{2}$$

It is also worth mentioning that $\alpha = 0 \mid b_e = 1$ is the event that SU transmits without collision.

Now, given that the SU is transmitting; i.e., $b_e = 1$, let $\boldsymbol{z} = \boldsymbol{n} + \alpha \boldsymbol{w}$. If the PU interference $\boldsymbol{w}$ follows a Gaussian distribution, we have $\boldsymbol{z}$ as a mixture of two Gaussian distributions, which is no longer Gaussian. In a more practical fading scenario where we have multiple primary transmitters using modulated signals, both $\boldsymbol{w}$ and $\boldsymbol{z}$ follow a GM distribution [32]. As such, in this work, we consider $\boldsymbol{w}$ as a mixture of p circularly symmetric complex Gaussian distributions, each with mean 0 and variance per real dimension $\sigma^2_{w,i}$, as follows:

$$f_W(\boldsymbol{w}) = \sum_{i=1}^{p} \gamma_i \mathcal{CN}(\boldsymbol{w}, 0, \sigma^2_{w,i}), \tag{3}$$

where γ_i is the mixing probability satisfying $\sum_{i=1}^{p} \gamma_i = 1$. As a result, $\boldsymbol{z}$ is a mixture of $(p + 1)$ Gaussian distributions as:

$$\begin{aligned} &f_z(z) = \\ &P(\alpha = 0 \mid b_e = 1)\mathcal{CN}(z, 0, \frac{N_0}{2}) + P(\alpha = 1 \mid b_e = 1) \sum_{i=1}^{p} \gamma_i \mathcal{CN}(z, 0, \sigma^2_{w,i} + \frac{N_0}{2}) \\ &= \sum_{i=1}^{p+1} \epsilon_i \mathcal{CN}(z, 0, \sigma_i^2), \end{aligned} \tag{4}$$

where $\epsilon_i = P(\alpha = 1 \mid b_e = 1)\gamma_i$, $1 \le i \le p$, and $\epsilon_{p+1} = P(\alpha = 0 \mid b_e = 1)$. In addition, $\sigma_i^2 = \sigma^2_{w,i} + \frac{N_0}{2}$ for $1 \le i \le p$, and $\sigma^2_{p+1} = \frac{N_0}{2}$. It can also be verified that $\boldsymbol{z}$ has mean of 0 and variance of $\sigma_z^2 = 2\sum_{i=1}^{p+1} \epsilon_i \sigma_i^2$.

2.2 Preliminaries on fundamental limits

The mutual information (MI) between the input $\boldsymbol{x}$ and the outputs $\boldsymbol{y}$ and b_e is given as

$$\mathcal{I}(\boldsymbol{x};\boldsymbol{y},b_e)=\mathcal{I}(\boldsymbol{x};b_e)+\mathcal{I}(\boldsymbol{x};\boldsymbol{y}|b_e)=\mathcal{P}(b_e=1)\,\mathcal{I}(\boldsymbol{x};\boldsymbol{y}|b_e=1), \quad (5)$$

where $P(b_e=1)=P_0\left(1-P_f\right)+P_1P_m$. For notational convenience, we define $\boldsymbol{y}_b$ as the channel output $\boldsymbol{y}$ conditioned on $b_e=1$, i.e., $\boldsymbol{y}_b \triangleq (\boldsymbol{y}|b_e=1)=\boldsymbol{h}_s\boldsymbol{x}+\boldsymbol{n}+\alpha\boldsymbol{w}$. Now, assume that an average power constraint $E(||\boldsymbol{x}||^2)\leq P_{in}$ is imposed on the input. The channel capacity is the supremum of the mutual information between the channel input and output over the set of all input distributions, and it can be expressed as

$$C=\sup_{\substack{f_X(\cdot)\\ E_X\left[\|\boldsymbol{x}\|^2\right]\leq P_{in}}} \mathcal{P}(b_e=1)\,\mathcal{I}(\boldsymbol{x};\boldsymbol{y}|b_e=1). \quad (6)$$

Equivalently, the channel capacity can be written as

$$C=\mathcal{P}(b_e=1)\cdot C_b. \quad (7)$$

and C_b is the channel capacity of the GM channel $\boldsymbol{y}_b=\boldsymbol{h}_s\boldsymbol{x}+\boldsymbol{n}+\alpha\boldsymbol{w}$. In the recent work in [43], it was shown that the optimal input for a GM channel has discrete amplitude with independent uniform phase. However, the explicit expression of the channel capacity still remains unknown.

On the other hand, energy efficiency is related to the channel capacity in the low-power regime and can be effectively measured by the energy per bit normalized by noise power. Generally, energy efficiency is maximized when energy-per-bit is minimized. For a given signal-to-noise ratio (SNR) defined as $\text{SNR}=\frac{E_X\left[\|\boldsymbol{x}\|^2\right]}{N_0}$, this normalized energy per bit is given as

$$\frac{E_b}{N_0}=\frac{\text{SNR}}{C\,(\text{SNR})}. \quad (8)$$

Then the minimum energy per bit required for reliable communication can be calculated as [45]

$$\frac{E_b}{N_{0\,\min}}=\lim_{\text{SNR}\to 0}\frac{\text{SNR}}{C\,(\text{SNR})}=\frac{\log_e 2}{\dot{C}(0)}. \quad (9)$$

Here, $\dot{C}(0)$ is the first derivative of the channel capacity at $\text{SNR}=0$. An input distribution is therefore considered first-order optimal if the first derivative of input/output mutual information at $\text{SNR}=0$ denoted as $\dot{I}(0)$ is equal to $\dot{C}(0)$.

Besides $\frac{E_b}{N_{0\,\min}}$, the wideband slope of the spectral efficiency as a function of E_b/N_0 calculated as $\subset\left(\frac{E_b}{N_0}\right)=C\,(\text{SNR})$ (in bits/sec/Hz) is another important fundamental limit in the low-power regime. This wideband slope (in bits/s/Hz per 3dB) is given by [45]

$$S_0=\frac{2\left(\dot{C}\,(0)\right)^2}{-\ddot{C}\,(0)}, \quad (10)$$

where $\dot{C}\,(0)$ and $\ddot{C}\,(0)$ are the first and second derivatives of the capacity at zero SNR. An input distribution is said to be second-order optimal if it is the first-order optimal and the second derivative of the input/output mutual information at $\text{SNR}=0$ denoted as $\ddot{I}(0)$ is equal to $\ddot{C}\,(0)$. By using a second-order optimal input signal, we can achieve a given rate at a given power with the minimum possible bandwidth.

3 Energy efficiency: minimum energy per bit and wideband slope

In this section, we establish in closed-form the minimum energy per bit $\frac{E_b}{N_{0\,\min}}$ and the wideband slope for the considered FD CR channel. Due to the lack of explicit expressions of both channel capacity and mutual information, we follow the approach in [45, 52] and rely on the relative entropy or Kullback-Leibler divergence without the need to identify the optimal input. Towards this end, in the following, we first calculate the two sensing metrics P_f and P_m taking into account the imperfectness of FD spectrum sensing. While spectrum sensing can be performed by different methods, such as energy detection, cyclostationary detection, and cooperative sensing, we simply consider energy detection [55–57]. The use of other methods is also possible, which simply leads to different sensing performances.

3.1 Sensing metrics: P_f and P_m

At a given sensing duration, whether or not the SU transmitter transmits data depends on the previous sensing result. In other words, unlike traditional cognitive radio systems, the statistics of the received energy at the SU which is used for sensing is not only a function of PU activity status, but also a function of SU activity. Consequently, based on the activity of PU and SU, there can be four different states at each sensing decision instant, $\left\{H_{ij}\right\}_{i,j\in\{0,1\}}$, where H_{ij} represent the state of $\hat{H}_i$ and H_j.

As such, to calculate P_f and P_m, we first need to examine two different cases of sensing: FD sensing where SU transmits and senses simultaneously, or HD sensing where only sensing is performed due to the negative result from the previous sensing decision. To describe these processes in further details, we can follow a similar approach in [26] and model the changes among the four states using two discrete-time Markov chains, as depicted in Fig. 2. Note that

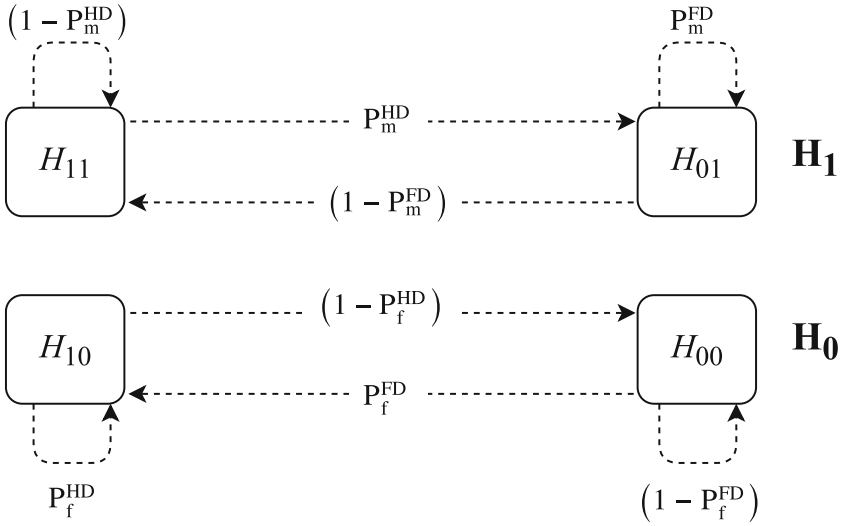

Fig. 2 SU state transitions for TS mode

in this structure we have assumed that PU activity changes sufficiently slow with respect to one SU time slot and we can neglect the rare time slots in which PU changes its state [24, 26, 58, 59].

We can then apply the steady-state analysis of Markov chains to calculate the probability of the system staying in each state $P\left(H_{ij}\right)$. It is then clear that $P_m = P\left(H_{01}\right)$, and $P_f = P\left(H_{10}\right)$.

In the case that PU is active, we have

$$\begin{cases} P\left(H_{11}\right) P_m^{HD} = P\left(H_{01}\right)\left(1-P_m^{FD}\right) \\ P\left(H_{01}\right) + P\left(H_{11}\right) = 1 \end{cases}. \tag{11}$$

So we can conclude that miss-detection probability is

$$P_m = P\left(H_{01}\right) = \frac{P_m^{HD}}{1 + P_m^{HD} - P_m^{FD}}. \tag{12}$$

Moreover, if the PU is not active, we have

$$\begin{cases} P\left(H_{10}\right)\left(1-P_f^{HD}\right) = P\left(H_{00}\right) P_f^{FD} \\ P\left(H_{00}\right) + P\left(H_{10}\right) = 1 \end{cases}, \tag{13}$$

and probability of false alarm is

$$P_f = P\left(H_{10}\right) = \frac{P_f^{FD}}{1 + P_f^{FD} - P_f^{HD}}. \tag{14}$$

To further evaluate P_m and P_f, let first consider the case of FD sensing. With energy detection-based sensing, the energy is calculated as

$$M = \frac{1}{f_s T_s} \sum_{n=1}^{f_s T_s} \left|\boldsymbol{y}\left(n\right)\right|^2, \tag{15}$$

where $\boldsymbol{y}\left(n\right)$ is the nth sampled signal at SU, and f_s and T_s are the sampling frequency and the sampling time, respectively. By comparing M with a pre-defined threshold ϵ, the false alarm and miss-detection probabilities are calculated as

$$P_f^{FD} = Pr\left[M > \epsilon \mid H_0\right], \tag{16}$$

and

$$P_m^{FD} = Pr\left[M < \epsilon \mid H_1\right]. \tag{17}$$

Under FD sensing, due the the presence of the data signal $\boldsymbol{x}$, the hypothesis tests of the the sampled signal $\boldsymbol{y}\left(n\right)$ are as follows:

$$\boldsymbol{y}\left(n\right) = \begin{cases} \boldsymbol{h}_{si}\left(n\right)\boldsymbol{x}\left(n\right) + \boldsymbol{n}\left(n\right), H_0 \\ \boldsymbol{w}\left(n\right) + \boldsymbol{h}_{si}\left(n\right)\boldsymbol{x}\left(n\right) + \boldsymbol{n}\left(n\right), H_1 \end{cases} \tag{18}$$

where $\boldsymbol{y}\left(n\right)$ is the n^{th} sample of the received signal at the SU receiver. In addition, and $\boldsymbol{h}_{si}\boldsymbol{x}\left(n\right)$ is the residual self-interference due to FD operation. In this work, we consider a realistic SI model in [60–62] in which $\boldsymbol{h}_{si}$ is modeled by a complex Gaussian distribution with zero mean and variance χ^2. This variance can be understood as the ratio of residual self-interference to the self-interference before suppression. Then we have the following proposition regarding the distribution of energy detected at SU receiver in FD mode, M_{FD}.

Proposition 1 *Under H_0, M_{FD} follows a Gaussian distribution with the following mean and variance*

$$E\left[M_{FD} \mid H_0\right] = \chi^2\sigma_s^2 + N_0, \tag{19}$$

and

$$Var\left[M_{FD} \mid H_0\right] = \frac{1}{f_s T_s}\left\{\chi^4\sigma_s^4\left(2\kappa_s - 1\right) + N_0^2 + 2\chi^2\sigma_s^2 N_0\right\}. \tag{20}$$

On the other hand, under H_1, M_{FD} is also Gaussian distributed, but with the following mean and variance

$$E\left[M_{FD} \mid H_1\right] = 2\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^2 + \chi^2\sigma_s^2 + N_0, \tag{21}$$

and

$$\begin{aligned} Var\left[M_{FD} \mid H_1\right] = \frac{1}{f_s T_s}\Bigg\{ & 8\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^4 + \chi^4\sigma_s^4\left(2\kappa_s - 1\right) \\ & + 4\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^2\left(-\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^2 + N_0\right) + 2\chi^2\sigma_s^2\left(2\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^2 + N_0\right) + N_0^2\Bigg\}. \end{aligned} \tag{22}$$

Given that, P_m^{FD} and P_f^{FD} can be calculated as:

$$P_m^{FD} = 1 - Q\left(\frac{\epsilon_{FD} - E\left[M_{FD} \mid H_1\right]}{\sqrt{Var\left[M_{FD} \mid H_1\right]}}\right), \tag{23}$$

and

$$P_f^{FD} = Q\left(\frac{\epsilon_{FD} - E\left[M_{FD} \mid H_0\right]}{\sqrt{Var\left[M_{FD} \mid H_0\right]}}\right). \tag{24}$$

The proof of this proposition is given in the Appendix. In the case of HD sensing, P_m^{HD} and P_f^{HD} can be obtained in a similar manner, but under the assumption that there is no residual SI, i.e., using $\chi = 0$ in Eqs. 23 and 24. Finally, substituting P_f^{HD}, P_m^{HD}, P_f^{HD}, and P_m^{HD} into Eqs. 12 and 14, we obtain P_m and P_f, respectively.

Given the above P_m P_f, we are ready to find the minimum energy per bit and the wideband slope, which is presented in the following subsection.

3.2 Minimum energy per bit and wideband slope

3.2.1 Minimum energy per bit

Based on the findings in [52], the first derivative of the capacity at zero SNR in nats can be calculated using the following equation:

$$\dot{C}\left(0\right) = N_0 \sup_{\boldsymbol{x}_0} \frac{D\left(f_{Y|X=\boldsymbol{x}_0}(\boldsymbol{y}) \parallel f_{Y|X=0}(\boldsymbol{y})\right)}{\|\boldsymbol{x}_0\|^2}, \tag{25}$$

where $f_{Y|X=\boldsymbol{x}_0}(\boldsymbol{y})$ and $f_{Y|X=0}(\boldsymbol{y})$ are the conditional PDFs of the output when the input $\boldsymbol{x} = \boldsymbol{x}_0$ and $\boldsymbol{x} = 0$, respectively and $D(\cdot)$ denotes the relative entropy between the two distributions. From earlier (9), it is clear that in order to calculate the minimum energy per bit for TS mode, it is enough to find the first derivative of capacity at zero SNR:

$$\frac{E_b}{N_{0\,\min}} = \lim_{\mathrm{SNR}\to 0}\frac{\mathrm{SNR}}{C\left(\mathrm{SNR}\right)} = \frac{\log_e 2}{\mathcal{P}\left(b_e = 1\right)\dot{C}_b\left(0\right)}. \tag{26}$$

For convenience, from Eq. 4, let $q = p + 1$. It can then be verified that

$$f_{Y_b|X=0}(\boldsymbol{y}) = \sum_{i=1}^{q}\epsilon_i\mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2), \tag{27}$$

and

$$f_{Y_b|X=x_0}(\boldsymbol{y}) = \sum_{i=1}^{q}\epsilon_i\mathcal{CN}(\boldsymbol{y}, \boldsymbol{h_s x}_0, \sigma_i^2). \tag{28}$$

Using the log-sum inequality, the relative entropy in Eq. 25 can be upper bounded as

$$\begin{aligned} & D\left(f_{Y_b|X=\mathbf{x_0}}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right) \\ & = \iint_{-\infty}^{\infty}\left(\sum_{i=1}^{q}\epsilon_i\mathcal{CN}(\boldsymbol{y}, \boldsymbol{h_s x}_0, \sigma_i^2)\right)\log\frac{\sum_{i=1}^{q}\epsilon_i\mathcal{CN}(\boldsymbol{y}, \boldsymbol{h_s x}_0, \sigma_i^2)}{\sum_{i=1}^{q}\epsilon_i\mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2)}d\mathbf{y} \\ & \le \iint_{-\infty}^{\infty}\sum_{i=1}^{q}\epsilon_i\mathcal{CN}(\boldsymbol{y}, \boldsymbol{h_s x}_0, \sigma_i^2).\log\frac{\mathcal{CN}(\boldsymbol{y}, \boldsymbol{h_s x}_0, \sigma_i^2)}{\mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2)}.d\mathbf{y} \\ & = \sum_{i=1}^{q}\epsilon_i D\left(\mathcal{CN}(\boldsymbol{y}, \boldsymbol{h_s x}_0, \sigma_i^2) \parallel \mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2)\right) \\ & = \sum_{i=1}^{q}\epsilon_i\frac{\|\boldsymbol{h_s x}_0\|^2}{2\sigma_i^2}\log e. \end{aligned} \tag{29}$$

The equality in Eq. 29 holds when $\frac{\epsilon_i\mathcal{CN}(\boldsymbol{y}, \boldsymbol{h_s x}_0, \sigma_i^2)}{\sum_{i=1}^{q}\epsilon_i\mathcal{CN}(\boldsymbol{y}, \boldsymbol{h_s x}_0, \sigma_i^2)} = \frac{\epsilon_i\mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2)}{\sum_{i=1}^{q}\epsilon_i\mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2)}$ for all i. In other words, the equality is achieved when $\|\boldsymbol{x}_0\| = 0$. Following from Eq. 25, the supremum is achieved as $\|\boldsymbol{x}_0\| \to 0$ and we have:

$$\begin{aligned} \dot{C}\left(0\right) & = \mathcal{P}\left(b_e = 1\right)\sigma_z^2 \sup_{\boldsymbol{x}_0}\frac{D\left(f_{Y_b|X=\boldsymbol{x}_0}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right)}{\|\boldsymbol{x}_0\|^2} \\ & = \mathcal{P}\left(b_e = 1\right)\sigma_z^2\|\boldsymbol{h_s}\|^2\sum_{1}^{q}\frac{\epsilon_i}{2\sigma_i^2} \\ & = \mathcal{P}\left(b_e = 1\right)\|\boldsymbol{h_s}\|^2\left(\sum_{i=1}^{q}2\epsilon_i\sigma_i^2\right)\left(\sum_{i=1}^{q}\frac{\epsilon_i}{2\sigma_i^2}\right). \end{aligned} \tag{30}$$

Therefore, from Eq. 9, the minimum bit energy can be calculated in closed-form as:

$$\frac{E_b}{N_{0\,\min}} = \frac{\log_e 2}{\|\boldsymbol{h_s}\|^2\mathcal{P}\left(b_e = 1\right)\left(\sum_{i=1}^{p+1}\epsilon_i\sigma_i^2\right)\left(\sum_{i=1}^{p+1}\frac{\epsilon_i}{\sigma_i^2}\right)}. \tag{31}$$

Note that in the case of traditional HD operation in TO mode, the minimum energy per bit can be obtained in a similar manner. Specifically, in this case, the calculations of $\mathcal{P}\left(b_e = 1\right)$ and $P\left(\alpha \mid b_e = 1\right)$ as well as the noise parameters in Eq. 4 are modified accordingly before (31) can be used to find $\frac{E_b}{N_{0\,\min}}$. For the brevity of the presentation, we omit these steps.

3.2.2 Wideband slope

In this section, we obtain the wideband slope in Eq. 10 in closed-form. Having found $\dot{C}\left(0\right)$ in previous part, we need

to formulate the second derivative of the capacity at zero SNR:

$$\ddot{C}(0) = \mathcal{P}(b_e = 1)\,\ddot{C}_b(0).$$

Given that the capacity-achieving input distribution of the channel $\boldsymbol{y}_b = \boldsymbol{h}_s\boldsymbol{x} + \boldsymbol{z}$ denoted as $P_{X_b^*}(\boldsymbol{x})$ is used, we have [45]

$$\ddot{C}_b(0) = 2\lim_{\mathrm{SNR}\to 0}\frac{C_b(\mathrm{SNR}) - \dot{C}_b(0)\,\mathrm{SNR}}{\mathrm{SNR}^2}. \tag{32}$$

As the capacity-achieving input distribution is the first-order optimal input, we can then simplify $\ddot{C}(0)$ as

$$\ddot{C}(0) = -2\mathcal{P}(b_e = 1)\lim_{\mathrm{SNR}\to 0}\frac{D\left(f_{Y_b^*}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right)}{\mathrm{SNR}^2}, \tag{33}$$

where $f_{Y_b^*}(\boldsymbol{y})$ is the output PDF when $P_{X_b^*}(\boldsymbol{x})$ is used.

As we are studying static channel, we can rewrite the channel as $\boldsymbol{y}_b = \boldsymbol{x}' + \boldsymbol{z}$ where $\boldsymbol{x}' = \boldsymbol{h}_s\boldsymbol{x}$. Using the polar coordinates, we rewrite the input and output as $\boldsymbol{x}' = re^{j\theta}$ and $\boldsymbol{y}_b = ve^{j\phi}$ where $r = \|\boldsymbol{h}_s\|\,\|\boldsymbol{x}\|$. It can then be verified that

$$D\left(f_{Y_b^*}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right) = D\left(f_{Y_b^*}(\boldsymbol{y}) \parallel f_{Y_b|X'=0}(\boldsymbol{y})\right) = D\left(f_{V^*}(v) \parallel f_{V|r=0}(v)\right) \tag{34}$$

where $f_{V^*}(v)$ is the PDF of the optimal output amplitude. In order to calculate the relative entropy in Eq. 34 we use the characterization of the optimal input, $P_{X_b^*}(\boldsymbol{x})$, having discrete amplitude and independently uniform phase that was shown recently in [43]. The PDF of the output v can then be written as[43]:

$$f_{V^*}(v) = \int_0^\infty \mathrm{K}(v, r)\, dF_{R^*}(r), \tag{35}$$

where $F_{R^*}(r)$ is the CDF of the optimal input amplitude and the Kernel $\mathrm{K}(v, r)$ is

$$\mathrm{K}(v, r) = \sum_{i=1}^{q}\frac{\epsilon_i}{\sigma_i^2}ve^{-\frac{v^2+r^2}{2\sigma_i^2}}\mathrm{I}_0\left(\frac{vr}{\sigma_i^2}\right). \tag{36}$$

Here, $I_0(\cdot)$ is the zero-order modified Bessel function. Since the amplitude of $P_{X_b^*}(\boldsymbol{x})$ consists of a finite number of mass points, let $\{p_k\}$ and $\{r_k\}$ be the probabilities and locations of these mass points, with $1 \le k \le K$. The PDF of the output amplitude under $P_{X_b^*}(\boldsymbol{x})$ can then be expressed as

$$f_{V^*}(v) = \sum_{k=1}^{K} p_k \mathrm{K}(v, r_k). \tag{37}$$

Moreover, we have

$$f_{V|r=0}(v) = \mathrm{K}(v, 0). \tag{38}$$

Then, using the log-sum inequality, we find the following upper bound on $D\left(f_{V^*}(v) \parallel f_{V|r=0}(v)\right)$:

$$D\left(f_{V^*}(v) \parallel f_{V|r=0}(v)\right) \le \sum_{k=1}^{K} p_k D\left(\mathrm{K}(v, r_k) \parallel \mathrm{K}(v, 0)\right). \tag{39}$$

It is straightforward to verify that the equality in Eq. 39 is obtained if and only if $r_k \to 0$ for $k = 1, 2, \cdots K$, which can be translated to $\mathrm{SNR} = \frac{\sum_{k=1}^{K} p_k r_k^2}{\sigma_z^2} \to 0$. Consequently, we have

$$\lim_{\mathrm{SNR}\to 0} D\left(f_{V^*}(v) \parallel f_{V|r=0}(v)\right) = \lim_{\mathrm{SNR}\to 0}\sum_{k=1}^{K} p_k D\left(\mathrm{K}(v, r_k) \parallel \mathrm{K}(v, 0)\right). \tag{40}$$

Without loss of generality, let's assume $r_1 < r_2 < \cdots < r_K$. As the function $D\left(\mathrm{K}(v, r) \parallel \mathrm{K}(v, 0)\right)$ is an increasing function of r, we obtain:

$$\lim_{\mathrm{SNR}\to 0} D\left(f_{V^*}(v) \parallel f_{V|r=0}(v)\right) \ge \lim_{\mathrm{SNR}\to 0}\sum_{k=1}^{K} p_k D\left(\mathrm{K}(v, r_1) \parallel \mathrm{K}(v, 0)\right). \tag{41}$$

Based on Eqs. 33, 34, and 41, we then conclude that

$$\ddot{C}(0) \le -2\mathcal{P}(b_e = 1)\lim_{\mathrm{SNR}\to 0}\frac{D\left(\mathrm{K}(v, r_1) \parallel \mathrm{K}(v, 0)\right)}{\mathrm{SNR}^2}. \tag{42}$$

Now, we need to show that the equality in Eq. 42 can be satisfied; i.e., the upper-bound is achievable. By using an input with one mass point amplitude and uniformly distributed phase we have $\mathrm{SNR} = \frac{r_1^2}{\|\boldsymbol{h}_s\|^2\sigma_z^2}$. For such input, it can easily be verified that the first derivative of the mutual information at zero SNR is equal to $\dot{C}(0)$ (calculated in Eq. 30), which means that this input is first-order optimal. Moreover, it can be seen that this input achieves the upper-bound in Eq. 42. Consequently, Eq. 42 is satisfied with equality.

Furthermore, by applying the log-sum inequality to the right-hand side of Eq. 42, we have

$$D\left(\mathrm{K}(v, r_1) \parallel \mathrm{K}(v, 0)\right) \le \sum_{i=1}^{q}\epsilon_i D\left(\mathrm{K}_i(v, r_1) \parallel \mathrm{K}_i(v, 0)\right), \tag{43}$$

where

$$\mathrm{K}_i(v, r) = \frac{1}{\sigma_i^2}ve^{-\frac{v^2+r^2}{2\sigma_i^2}}\mathrm{I}_0\left(\frac{vr}{\sigma_i^2}\right). \tag{44}$$

It is clear that the equality in Eq. 43 is achieved when $r_1 = 0$. Thus, $\ddot{C}(0)$ can be formulated as

$$\ddot{C}(0) = -2\mathcal{P}(b_e=1)\left(\sigma_z^2\right)^2 \|\boldsymbol{h}_s\|^4 \times \lim_{r\to 0}\frac{1}{r^4}\sum_{i=1}^{q}\frac{\epsilon_i}{\sigma_i^2} * \int_0^\infty \left\{ v e^{-\frac{v^2+r^2}{2\sigma_i^2}} \mathrm{I}_0\left(\frac{vr}{\sigma_i^2}\right) \times \log\left[e^{-\frac{r^2}{2\sigma_i^2}} \mathrm{I}_0\left(\frac{vr}{\sigma_i^2}\right)\right]\right\} dv. \quad (45)$$

Because the integrand in Eq. 45 and its derivatives are continuous in r for $0 \le r, v < \infty$, we use Leibniz integral rule and L'Hospital's rule to simplify (45). In particular, by applying L'Hospital's rule for four times, we obtain the following:

$$\begin{aligned}\ddot{C}(0) &= -2\mathcal{P}(b_e=1)\left(\sigma_z^2\right)^2 \|\boldsymbol{h}_s\|^4 \sum_{i=1}^{q}\frac{\epsilon_i}{\sigma_i^2}\int_0^\infty e^{-\frac{v^2}{2\sigma_i^2}}\left(\frac{3v^5}{64\sigma_i^8} - \frac{v^3}{4\sigma_i^6} + \frac{v}{4\sigma_i^4}\right)dv \\ &= -2\mathcal{P}(b_e=1)\|\boldsymbol{h}_s\|^4\left(\sum_{i=1}^{q} 2\epsilon_i\sigma_i^2\right)^2 \sum_{i=1}^{q}\frac{\epsilon_i}{8\sigma_i^4}. \end{aligned} \quad (46)$$

Then from Eq. 10, we can find the wideband slope for TS mode as

$$S_0 = 2\mathcal{P}(b_e=1)\frac{\left[\sum_1^{p+1}\frac{\epsilon_i}{\sigma_i^2}\right]^2}{\sum_1^{p+1}\frac{\epsilon_i}{\sigma_i^4}}. \quad (47)$$

Note that the expression in Eq. 47 is in closed-form, and the wideband slope can be easily calculated. As similar to the calculation of $\frac{E_b}{N_0}_{\min}$, the wideband slope for the traditional HD in TO mode can also be obtained by using modified $\mathcal{P}(b_e = 1)$.

4 Optimal signaling schemes for FD CR under imperfect spectrum sensing

Now that we have found both the minimum energy per bit and the wideband slope in closed-form, it is of interest to analyze the first-order and second-order optimal signaling schemes.

As we discussed in previous section, an input distribution is considered first-order optimal if $\dot{\mathcal{I}}(0)$ is equal to $\dot{C}(0)$ and a first order optimal input distribution is said to be second-order optimal if $\ddot{\mathcal{I}}(0)$ is equal to $\ddot{C}(0)$. Therefore in the next following parts of this section, we derive a necessary condition for an input to be first-order optimal and examine several practical inputs to investigate their second-order optimality.

4.1 A necessary condition for first-order optimality

In this part, we find a necessary condition for first-order optimality of an input. As we stated before, an input is considered to be first-order optimal if

$$\lim_{\mathrm{SNR}\to 0}\frac{\mathcal{I}(\boldsymbol{x},\boldsymbol{y})}{\mathrm{SNR}} = \dot{C}(0). \quad (48)$$

For a specific input distribution, the mutual information between the output and the input is:

$$\mathcal{I}(\boldsymbol{x},\boldsymbol{y}) = \mathcal{P}(b_e=1)\,\mathcal{I}\left(\boldsymbol{x},\boldsymbol{y}_b\right).$$

By writing the mutual information $\mathcal{I}(\boldsymbol{x},\boldsymbol{y})$ in the canonical form, we then obtain:

$$\mathcal{I}(\boldsymbol{x},\boldsymbol{y}) = \mathcal{P}(b_e=1)\left\{E_X\left[D\left(f_{Y_b|X=\boldsymbol{x}}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right)\right] - D(f_{Y_b}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y}))\right\}. \quad (49)$$

From previous section, we have $D\left(f_{Y_b|X=\boldsymbol{x}}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right) \le \|\boldsymbol{h}_s\|^2 \sum_1^q \frac{\epsilon_i.\|\boldsymbol{x}\|^2}{2\sigma_i^2}$. Therefore,

$$\begin{aligned} E_X\left[D\left(f_{Y_b|X=\boldsymbol{x}}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right)\right] &\le \|\boldsymbol{h}_s\|^2 E_X\left[\sum_{i=1}^{q}\frac{\epsilon_i.\|\boldsymbol{x}\|^2}{2\sigma_i^2}\right] \\ &= \|\boldsymbol{h}_s\|^2 E_X\left[\|\boldsymbol{x}\|^2\right]\sum_{i=1}^{q}\frac{\epsilon_i}{2\sigma_i^2}. \end{aligned} \quad (50)$$

When $\mathrm{SNR} \to 0$, $E_X\left[\|\boldsymbol{x}\|^2\right]$ approaches zero. As a result, the input distribution behaves like a Dirac Delta function, i.e., $f_X(\boldsymbol{x})$ approaches $\delta(\boldsymbol{x} - E_X[\boldsymbol{x}])$. Hence,

$$\begin{aligned} &\lim_{\mathrm{SNR}\to 0}\frac{\mathcal{P}(b_e=1)\,E_X\left[D\left(f_{Y_b|X=\boldsymbol{x}}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right)\right]}{\mathrm{SNR}} \\ &= \lim_{\mathrm{SNR}\to 0}\sigma_z^2\frac{\mathcal{P}(b_e=1)\,E_X\left[D\left(f_{Y_b|X=\boldsymbol{x}}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right)\right]}{E_X\left[\|\boldsymbol{x}\|^2\right]} \\ &= \dot{C}(0). \end{aligned} \quad (51)$$

So the first term on the right-hand side of Eq. 49 approaches $\dot{C}(0)$ when $\mathrm{SNR} \to 0$. As such, we now need to find a condition from which the term $D(f_{Y_b}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y}))$ in Eq. 49 is $o(\mathrm{SNR})$. Therefore, this condition guarantees the first order optimality of an input. Toward this end, we re-write the output distribution as follows:

$$f_{Y_b}(\boldsymbol{y}) = \sum_{i=1}^{q}\epsilon_i h_i(\boldsymbol{y}). \quad (52)$$

Here, $h_i(\boldsymbol{y})$ is the convolution of $\boldsymbol{h_s x}$ and i^{th} term of noise distribution, f_Z. It is then clear that $h_i(\boldsymbol{y})$ has mean $\boldsymbol{h_s} E_X[\boldsymbol{x}]$ and covariance matrix $\sigma_i^2\boldsymbol{I}_{2\times 2} + \Sigma_{h_s x}$, where $\Sigma_{h_s x}$ is the covariance matrix of the distribution of $\boldsymbol{h_s x}$. We then have the following inequality:

$$D\left(f_{Y_b}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right) \le \sum_{i=1}^{q}\epsilon_i D\left(h_i(\boldsymbol{y}) \parallel \mathcal{CN}(\boldsymbol{y},0,\sigma_i^2)\right). \quad (53)$$

Each term on the right-hand side of Eq. 53 can be further expanded as:

$$D\left(h_i(\boldsymbol{y}) \parallel \mathcal{CN}(\boldsymbol{y},0,\sigma_i^2)\right) = D\left(h_i(\boldsymbol{y}) \parallel \phi_{h_i}(\boldsymbol{y})\right) + D\left(\phi_{h_i}(\boldsymbol{y}) \parallel \mathcal{CN}(\boldsymbol{y},0,\sigma_i^2)\right), \tag{54}$$

where $\phi_{h_i}(\boldsymbol{y})$ is a Gaussian distribution with the same mean and covariance matrix as $h_i(\boldsymbol{y})$. In [63], it is proved that

$$D\left(h_i(\boldsymbol{y}) \parallel \phi_{h_i}(\boldsymbol{y})\right) = o\left(\text{SNR}\right). \tag{55}$$

As a result, we need to focus only on the second term $D\left(\phi_{h_i}(\boldsymbol{y}) \parallel \mathcal{CN}(\boldsymbol{y},0,\sigma_i^2)\right)$. For that, we have the following lemma:

Lemma 1 *For an input* $\boldsymbol{x}$*, if*

$$\lim_{SNR\to 0} \frac{\|E_X[\boldsymbol{x}]\|^2}{E_X\left[\|\boldsymbol{x}\|^2\right]} = 0, \tag{56}$$

then $D\left(\phi_{h_i}(\boldsymbol{y}) \parallel \mathcal{CN}(\boldsymbol{y},0,\sigma_i^2)\right) = o\left(SNR\right)$

Proof We have

$$D\left(\phi_{h_i}(\boldsymbol{y}) \parallel \mathcal{CN}(\boldsymbol{y},0,\sigma_i^2)\right) = \frac{\|\boldsymbol{h_s}\|^2 \|E_X[\boldsymbol{x}]\|^2}{\sigma_i^2} + \log\left(\det\left(\sigma_i^2 I\right)\right)$$

$$- \log\left(\det\left(\sigma_i^2 I + \Sigma_{h_s x}\right)\right) + Tr\left(\sigma_i^{-2} I\left(\sigma_i^2 I + \Sigma_{h_s x}\right) - I\right) \tag{57}$$

$$= -\log\left(\det\left(I + \sigma_i^{-2}\Sigma_{h_s x}\right)\right) + Tr\left(\sigma_i^{-2}\Sigma_{h_s x}\right) + \frac{\|\boldsymbol{h_s}\|^2 \|E_X[\boldsymbol{x}]\|^2}{\sigma_i^2} \tag{58}$$

$$\leq \frac{1}{2\sigma_i^4} Tr(\Sigma_{h_s x}^2) + \frac{\|\boldsymbol{h_s}\|^2 \|E_X[\boldsymbol{x}]\|^2}{\sigma_i^2} \tag{59}$$

$$\leq \frac{1}{2\sigma_i^4} Tr^2(\Sigma_{hx}) + \frac{\|\boldsymbol{h_s}\|^2 \|E_X[\boldsymbol{x}]\|^2}{\sigma_i^2} \tag{60}$$

$$= \frac{\|\boldsymbol{h_s}\|^2}{2\sigma_i^4} Tr^2(\Sigma_x) + \frac{\|\boldsymbol{h_s}\|^2 \|E_X[\boldsymbol{x}]\|^2}{\sigma_i^2}. \tag{61}$$

Note that in Eqs. 57 and 58, we have used the following:

$$D\left(\mathcal{CN}(\boldsymbol{y},\mathbf{m_0},\Sigma_1) \parallel \mathcal{CN}(\boldsymbol{y},\mathbf{m_1},\Sigma_0)\right) =$$
$$\log\left(\det\Sigma_0\right) - \log\left(\det\Sigma_1\right) + (\boldsymbol{m}_1 - \boldsymbol{m}_0)^T \Sigma_0^{-1}(\boldsymbol{m}_1 - \boldsymbol{m}_0) + Tr\left(\Sigma_0^{-1}\Sigma_1 - I\right) \tag{62}$$

and

$$\log\left(\det(I + A)\right) \geq Tr(A) - 0.5Tr\left(A^2\right). \tag{63}$$

for a non-negative matrix A. Equation 60 also comes from the fact that $Tr\left(\Sigma_{h_s x}\right) = \|\boldsymbol{h_s}\|^2 Tr\left(\Sigma_x\right)$ because $\boldsymbol{h_s}$ is a known value. Then combining Eqs. 60 and 53, we have:

$$D\left(f_{Y_b}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right) \leq \|\boldsymbol{h_s}\|^2 \left(Tr^2(\Sigma_X)\sum_1^q \frac{\epsilon_i}{2\sigma_i^4} + \|E_X[\boldsymbol{x}]\|^2 \sum_1^q \frac{\epsilon_i}{\sigma_i^2}\right). \tag{64}$$

As $E_X\left[\|\boldsymbol{x}\|^2\right] = Tr\left(\Sigma_x\right)$, it is then clear that the condition in Eq. 56 guarantees $D\left(f_{Y_b}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right) = o\left(\text{SNR}\right)$. □

Based on Lemma 1, it is clear that Eq. 56 is a necessary condition for the first order optimality of an input.

4.2 Second-order optimal signaling

In previous part, we found a necessary condition for first-order optimality. Based on this condition, it is clear that a wide range of inputs are first-order optimal. Thus, It is of interest to investigate simple distributions that are second-order optimal. Therefore, we are consider several practical input signaling schemes to analyze their second-order optimality. As such we consider: i) An input having a single mass point amplitude and independently uniform phase; ii) a Gaussian input; iii) BPSK; and iv) QPSK.

4.2.1 An input with a single mass point amplitude and independently uniform phase

Using numerical results, it was shown in [43] that this input is capacity-achieving for Gaussian-Mixture channels at sufficiently low SNRs. Besides, it is straightforward to verify that the equality in Eq. 42 is achieved by using an input with a single mass point amplitude and independently uniform phase. Hence, the results on $E_b/N_{0\min}$ and the wideband slope, analytically confirms that this input is optimal in the low-SNR regime.

4.2.2 A Gaussian input

We consider a Gaussian input with mean 0 and variance per dimension σ_X^2, which is first-order optimal. The output PDF can then be written as:

$$f_{Y_b}(\boldsymbol{y}) = \sum_{i=1}^{q} \epsilon_i \mathcal{CN}(\boldsymbol{y},0,\sigma_i^2 + \|\boldsymbol{h_s}\|^2\sigma_X^2). \tag{65}$$

Using Log-Sum inequality, we then have:

$$D\left(f_{Y_b}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right) \leq \sum_{i=1}^{q} \epsilon_i D\left(\mathcal{CN}(\boldsymbol{y},0,\sigma_i^2 + \|\boldsymbol{h_s}\|^2\sigma_X^2) \parallel \mathcal{CN}(\boldsymbol{y},0,\sigma_i^2)\right)$$
$$= \sum_{i=1}^{q} \epsilon_i \left(\log\frac{\sigma_i^2}{\sigma_i^2 + \|\boldsymbol{h_s}\|^2\sigma_X^2} + \frac{\|\boldsymbol{h_s}\|^2\sigma_X^2}{\sigma_i^2}\right) \tag{66}$$

with equality iff $\sigma_X^2 = 0$. It then follows that:

$$\ddot{\mathcal{I}}(0) = -2\mathcal{P}(b_e = 1)\lim_{\text{SNR}\to 0}\frac{D\left(f_{Y_b}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right)}{\text{SNR}^2}$$
$$\geq -2\mathcal{P}(b_e = 1)\lim_{\text{SNR}\to 0}\frac{\sum_{i=1}^{q}\epsilon_i\left(\log\frac{\sigma_i^2}{\sigma_i^2 + \|\boldsymbol{h_s}\|^2\sigma_X^2} + \frac{\|\boldsymbol{h_s}\|^2\sigma_X^2}{\sigma_i^2}\right)}{\text{SNR}^2}. \tag{67}$$

As SNR $= \frac{E\left[\|\boldsymbol{x}\|^2\right]}{\sigma_z^2} = \frac{2\sigma_x^2}{\sigma_z^2}$, the equality in Eq. 67 is achieved when SNR $\to 0$. So, we obtain:

$$\ddot{\mathcal{I}}(0) = -2\mathcal{P}(b_e = 1)\,\sigma_z^4\,\|\boldsymbol{h_s}\|^4 \sum_{i=1}^{q} \frac{\epsilon_i}{8\sigma_i^4}. \tag{68}$$

As a result, S_0 (Gaussian) $= \frac{2\left[\dot{C}(0)\right]^2}{-\ddot{\mathcal{I}}(0)} = \frac{2\mathcal{P}(b_e=1)\left[\sum_{i=1}^{q}\frac{\epsilon_i}{\sigma_i^2}\right]^2}{\sum_{i=1}^{q}\frac{\epsilon_i}{\sigma_i^4}}$.
It means that the Gaussian input is second-order optimal.

4.2.3 BPSK

Now. let's consider the BPSK modulation with the following PDF:

$$P_X(\boldsymbol{x}) = \frac{1}{2}\delta_A + \frac{1}{2}\delta_{-A}. \tag{69}$$

The output PDF can then be expressed as:

$$\begin{aligned} f_{Y_b}(\boldsymbol{y}) &= \sum_{i=1}^{t} \frac{\epsilon_i}{2}\left(\mathcal{CN}\left(\boldsymbol{y}, \boldsymbol{h_s}A, \sigma_i^2\right) + \mathcal{CN}\left(\boldsymbol{y}, -\boldsymbol{h_s}A, \sigma_i^2\right)\right) \\ &= \sum_{i=1}^{q} \epsilon_i h_i(\boldsymbol{y}), \end{aligned} \tag{70}$$

where $h_i(\boldsymbol{y})$ is the convolution of distribution of $\boldsymbol{h_s x}$ and i^{th} element of f_Z and can be written as:

$$\begin{aligned} h_i(\boldsymbol{y}) &= \frac{1}{4\pi\sigma_i^2}\left[\exp\left(-\frac{|\boldsymbol{y}-\boldsymbol{h_s}A|^2}{2\sigma_i^2}\right) + \exp\left(-\frac{|\boldsymbol{y}+\boldsymbol{h_s}A|^2}{2\sigma_i^2}\right)\right] \\ &= \frac{\exp\left(\frac{-\|\boldsymbol{h_s}\|^2A^2}{2\sigma_i^2}\right)}{2\pi\sigma_i^2}\exp\left(\frac{-|\boldsymbol{y}|^2}{2\sigma_i^2}\right)\cosh\left(\frac{A\Re\left(\boldsymbol{h_s}^*\boldsymbol{y}\right)}{\sigma_i^2}\right) \\ &= \exp\left(\frac{-\|\boldsymbol{h_s}\|^2A^2}{2\sigma_i^2}\right)\cosh\left(\frac{A\Re\left(\boldsymbol{h_s}^*\boldsymbol{y}\right)}{\sigma_i^2}\right)\mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2). \end{aligned} \tag{71}$$

Based on log-sum inequality, we obtain

$$D\left(f_{Y_b}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right) \le \sum_{1}^{q}\epsilon_i D\left(h_i \parallel \mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2)\right), \tag{72}$$

with equality iff $A = 0$. Moreover, it is easy to verify that

$$D\left(h_i(\boldsymbol{y}) \parallel \mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2)\right) = E\left\{\phi_i(\boldsymbol{y})\log\phi_i(\boldsymbol{y})\right\}, \tag{73}$$

where

$$\phi_i(\boldsymbol{y}) = \exp\left(\frac{-\|\boldsymbol{h_s}\|^2A^2}{2\sigma_i^2}\right)\cosh\left(\frac{A\Re\left(\boldsymbol{h_s}^*\boldsymbol{y}\right)}{\sigma_i^2}\right), \tag{74}$$

and the expectation is done over a Gaussian distribution with mean zero and variance $2\sigma_i^2$. due to the fact that $\frac{A\Re(\boldsymbol{h_s}^*\boldsymbol{y})}{\sigma_i^2}$ follows a Gaussian distribution with variance $\frac{\|\boldsymbol{h_s}\|^2A^2}{\sigma_i^2}$, we have

$$D\left(h_i(\boldsymbol{y}) \parallel \mathcal{CN}(\boldsymbol{y}, 0, \sigma_i^2)\right) = \left(\frac{\|\boldsymbol{h_s}\|^2A^2}{2\sigma_i^2}\right)^2 + o\left(\left(\frac{\|\boldsymbol{h_s}\|^2A^2}{2\sigma_i^2}\right)^2\right). \tag{75}$$

Since SNR $= \frac{A^2}{\sigma_z^2}$, we then obtain:

$$\begin{aligned} \ddot{\mathcal{I}}(0) &= -2\mathcal{P}(b_e = 1)\lim_{\text{SNR}\to 0}\frac{D\left(f_{Y_b}(\boldsymbol{y}) \parallel f_{Y_b|X=0}(\boldsymbol{y})\right)}{\text{SNR}^2} \\ &\ge -2\mathcal{P}(b_e = 1)(N_0)^2\lim_{A\to 0}\frac{\sum_1^q\epsilon_i\left[\left(\frac{\|\boldsymbol{h_s}\|^2A^2}{2\sigma_i^2}\right)^2 + o\left(\left(\frac{\|\boldsymbol{h_s}\|^2A^2}{2\sigma_i^2}\right)^2\right)\right]}{A^4} \end{aligned} \tag{76}$$

with equality iff $A = 0$. As a result,

$$\ddot{\mathcal{I}}(0) = -2\mathcal{P}(b_e = 1)\,\sigma_z^4\,\|\boldsymbol{h_s}\|^4\sum_{1}^{q}\frac{\epsilon_i}{4\sigma_i^4}, \tag{77}$$

and the wideband slope can be found as

$$S_0(\text{BPSK}) = \frac{2\left[\dot{C}(0)\right]^2}{-\ddot{\mathcal{I}}(0)} = \frac{\mathcal{P}(b_e = 1)\left[\sum_1^q\frac{\epsilon_i}{\sigma_i^2}\right]^2}{\sum_1^q\frac{\epsilon_i}{\sigma_i^4}}. \tag{78}$$

The wideband slope of BPSK is only half of the optimal one, which means BPSK input is not second-order optimal

4.2.4 QPSK

It is straightforward to verify that

$$C_{QPSK}(\text{SNR}) = 2C_{BPSK}\left(\frac{\text{SNR}}{2}\right). \tag{79}$$

Moreover, we have

$$\left(\frac{E_b}{N_0}\right)_{QPSK} = \frac{\text{SNR}}{C_{QPSK}(\text{SNR})}. \tag{80}$$

Hence, for the same energy per bit, i.e., for

$$\left(\frac{E_b}{N_0}\right)_{BPSK} = \left(\frac{E_b}{N_0}\right)_{QPSK}. \tag{81}$$

QPSK clearly achieves twice the spectral efficiency of BPSK, and, consequently, twice the wideband slope. As a result, QPSK is second-order optimal.

5 Numerical result

In this section, numerical results are provided to confirm the analysis on the sensing performance and the spectral efficiency of full duplex cognitive radio systems. Unless we state otherwise, the following setup is adopted. The

sampling frequency is chosen at $f_s = 5$MHz, and we assume $\frac{\sigma_s^2}{N_0} = 20$dB, and $P_0 = P_1 = 0.5$. Moreover, for the primary interference on secondary network, we consider a 3-term Gaussian mixture noise channel with $\epsilon_1 = 0.8$, $\sigma_1^2 = 0.1$, $\epsilon_2 = 0.1$, $\sigma_2^2 = 1$, $\epsilon_3 = 0.1$, $\sigma_3^2 = 10$.

5.1 Spectrum sensing

Let first examine the sensing performance in the considered full duplex cognitive radio system under practical self-interference. For simplicity, we assume that the primary signal is complex PSK modulated. Similar results can be obtained for any modulation scheme.

For the sensing performance, we use the receiver operating characteristic (ROC) curve, which is a helpful tool to visualize the trade-off between the probability of false alarm and the probability of miss-detection. The ROC curves for TS operation for different values of χ are plotted in Fig. 3. For comparison, the corresponding ROC curve for traditional TO sensing is also provided in Fig. 3 for different $\frac{T_s}{T}$. First, it is clear that a larger χ results in a poorer performance of TS-based sensing. It can also be observed from Fig. 3 that the sensing performance of TS mode is better than TO-based sensing when the residual self interference is sufficiently low, e.g., it depends on the performance of self-interference cancellation.

5.2 $\frac{E_b}{N_0\,min}$ and wideband slope

Now, let examine the spectral efficiency of FD CR in terms of $\frac{E_b}{N_0\,min}$ and the wideband slope to demonstrate the benefits offered by FD CR. It should also be noted that these two fundamental metrics have been obtained in closed-form, and they can be easily calculated without the need of simulations.

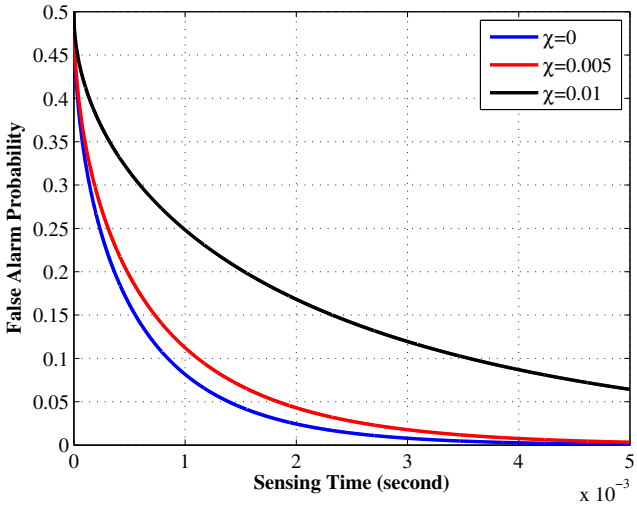

Fig. 3 ROC curves for TS and HD sensing for different values of χ and $\frac{T_s}{T}$

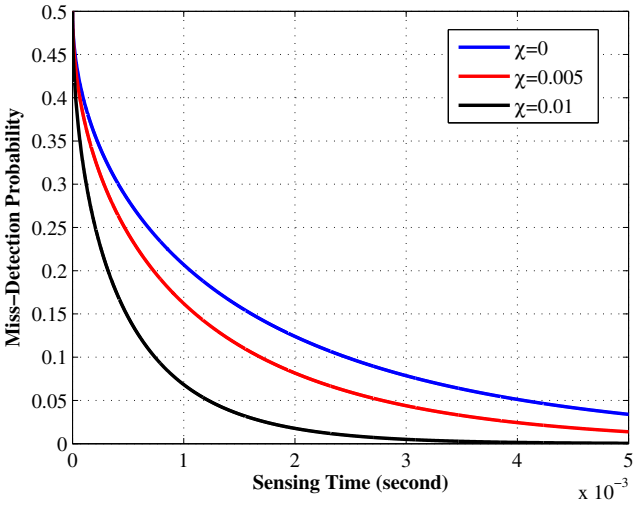

Fig. 4 $\frac{E_b}{N_0\,min}$ vs frame length T achieved by TS operation with different values of χ

Figure 4 first shows $\frac{E_b}{N_0\,min}$ of the TS-based system achieved with different values of frame time T and χ. It is clearly that the quality of self-interference cancellation plays an important role on $\frac{E_b}{N_0\,min}$. Interestingly, $\frac{E_b}{N_0\,min}$ with $\chi = 0.01$ is almost identical to the case of perfect self-interference cancellation with $\chi = 0$. Furthermore, it can be observed from Fig. 4 that for a given χ, when T increases, $\frac{E_b}{N_0\,min}$ is getting smaller. However, the minimum energy per bit approaches a constant at a sufficiently large T. This behavior can be confirmed analytically from Eq. 31 where it can be shown that the quality of sensing cannot be improved further by increasing T. As a result, $\frac{E_b}{N_0\,min}$ asymptotically approaches a fixed value when T increases.

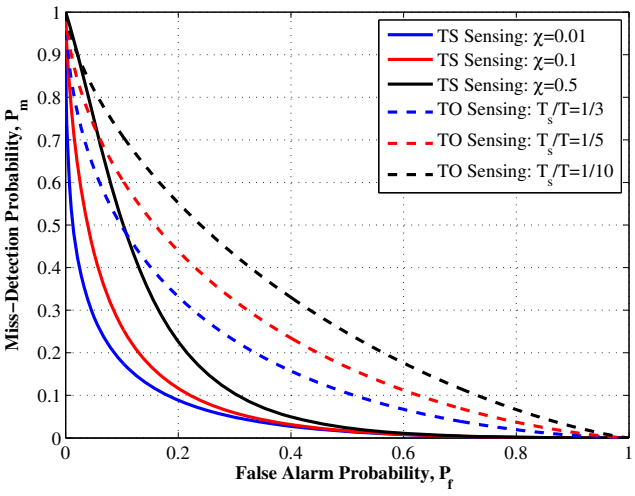

Fig. 5 $\frac{E_b}{N_0\,min}$ vs frame length T achieved by TS operation with $\chi = 0.01$ and TO operation. For TO, $T_s = \frac{10^{-3}}{3}$ (seconds)

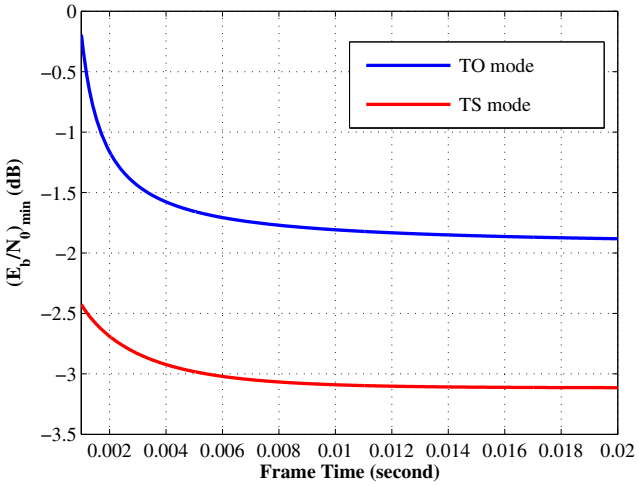

Fig. 6 Wideband slope vs frame length T achieved by TS operation with different values of χ

Figure 5 compares $\frac{E_b}{N_0}_{min}$ of TO and TS modes using practical $\chi = 0.01$ over a wide range of frame length T. For TO mode, the sensing time is fixed at $T_s = \frac{10^{-3}}{3}$ seconds. The benefits of FD over HD can be clearly observed where the TS-based system achieves significantly better $\frac{E_b}{N_0}_{min}$.

Similar results in terms of the wideband slope are also obtained. In particular, Fig. 6 plots the wideband slope achieved by the TS-based system using different values of χ, while the wideband slope comparison between the TS-based system with $\chi = 0.01$ and TO-based system is shown in Fig. 7. Note that for the TO-based system, T_s is set at $\frac{10^{-3}}{3}$

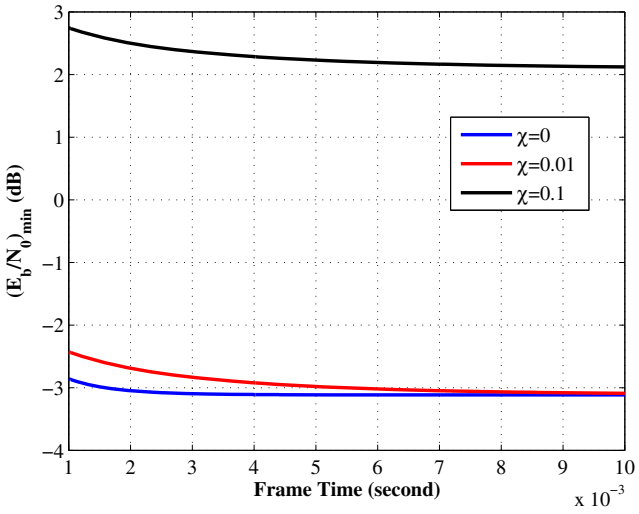

Fig. 7 Wideband slope vs frame length T achieved by TS operation with $\chi = 0.01$ and TO operation. For TO, $T_s = \frac{10^{-3}}{3}$ (seconds)

(seconds). It is clear from Fig. 7 that TS operation achieves a significantly higher wideband slope.

6 Conclusion

In this work, we have proposed an effective way to evaluate the energy efficiency of full duplex (FD) cognitive radio (CR) in the low-power regime under practical self-interference cancellation and imperfect spectrum sensing. Specifically, the two fundamental limits of minimum energy per bit and wideband slope of the considered FD CR systems were established in closed-form. From these limits, practical yet optimal signaling schemes at the low-SNR regime were identified. The benefits in terms of both minimum energy per bit and wideband slope of FD over HD were also clearly demonstrated.

Appendix

Proposition 2 *Using the Central Limit Theorem, for a large number of samples (f_sT_s), the distribution of M_{FD} given H_0 can be approximated by Gaussian distribution with the following mean and variance*

$$E\left[M_{FD} \mid H_0\right] = \chi^2\sigma_s^2 + N_0, \tag{82}$$

$$Var\left[M_{FD} \mid H_0\right] = \frac{1}{f_sT_s}\left\{\chi^4\sigma_s^4\left(2\kappa_s - 1\right) + N_0^2 + 2\chi^2\sigma_s^2N_0\right\}. \tag{83}$$

Proof

$$\begin{aligned} &E\left[M_{FD} \mid H_0\right] = E\left[\left|\boldsymbol{y}_{FD}(n)\right|^2 \mid H_0\right] = E\left[\left|\boldsymbol{h}_{si}(n)\boldsymbol{x}(n) + \boldsymbol{n}(n)\right|^2\right] \\ &= E\left[\left|\boldsymbol{h}_{si}(n)\right|^2\left|\boldsymbol{x}(n)\right|^2 + \left|\boldsymbol{n}(n)\right|^2 + \boldsymbol{h}_{si}(n)\boldsymbol{x}(n)\boldsymbol{n}^*(n) + \boldsymbol{h}_{si}^*(n)\boldsymbol{x}_s^*(n)\boldsymbol{n}(n)\right] \\ &= \chi^2\sigma_s^2 + N_0. \end{aligned} \tag{84}$$

Moreover,

$$\begin{aligned} Var\left[M_{FD} \mid H_0\right] &= \frac{1}{f_sT_s}Var\left[\left|\boldsymbol{y}_{FD}(n)\right|^2 \mid H_0\right] \\ &= \frac{1}{f_sT_s}\left\{E\left[\left|\boldsymbol{y}_{FD}(n)\right|^4 \mid H_0\right] - \left(E\left[\left|\boldsymbol{y}_{FD}(n)\right|^2 \mid H_0\right]\right)^2\right\}. \end{aligned} \tag{85}$$

We already have $E\left[\left|\boldsymbol{y}_{FD}(n)\right|^2 \mid H_0\right]$ in Eq. 84. So we only need to calculate $E\left[\left|\boldsymbol{y}_{FD}(n)\right|^4 \mid H_0\right]$.

$$\begin{aligned} &E\left[\left|\boldsymbol{y}_{FD}(n)\right|^4 \mid H_0\right] = E\left[\left|\boldsymbol{h}_{si}(n)\boldsymbol{x}(n) + \boldsymbol{n}(n)\right|^4\right] \\ &= E\left[\left|\left|\boldsymbol{h}_{si}(n)\right|^2\left|\boldsymbol{x}(n)\right|^2 + \left|\boldsymbol{n}(n)\right|^2 + \boldsymbol{h}_{si}(n)\boldsymbol{x}(n)\boldsymbol{n}^*(n) + \boldsymbol{h}_{si}^*(n)\boldsymbol{x}_s^*(n)\boldsymbol{n}(n)\right|^2\right] \\ &= E\left[\left|\boldsymbol{h}_{si}(n)\right|^4\left|\boldsymbol{x}(n)\right|^4 + \left|\boldsymbol{n}(n)\right|^4 + \left|\boldsymbol{h}_{si}(n)\boldsymbol{x}(n)\boldsymbol{n}^*(n)\right|^2\right. \\ &\qquad \left. + \left|\boldsymbol{h}_{si}^*(n)\boldsymbol{x}_s^*(n)\boldsymbol{n}(n)\right|^2 + 2\left|\boldsymbol{h}_{si}(n)\right|^2\left|\boldsymbol{x}(n)\right|^2\left|\boldsymbol{n}(n)\right|^2\right] \\ &= 2\chi^4\kappa_s\sigma_s^4 + 2N_0^2 + 4\chi^2\sigma_s^2N_0. \end{aligned} \tag{86}$$

In the above expression we have used the fact that for circularly symmetric complex random varibale X with mean zero we have

$$E\left[X^2\right] = E\left[(X_r + jX_i)^2\right] = E\left[|X_r|^2\right] - E\left[|X_i|^2\right] + 2jE\left[X_r X_i\right] = 0 \quad (87)$$

By using Eqs. 84, 85, and 86, Eq. 83 is proved. □

Proposition 3 *Using the Central Limit Theorem, for the large number of samples ($f_s T_s$), the distribution of M_{FD} given H_1 can be approximated by Gaussian distribution with the following mean and variance*

$$E\left[M_{FD} \mid H_1\right] = 2\sum_{i=1}^{p} \gamma_i \sigma_{w,i}^2 + \chi^2 \sigma_s^2 + N_0, \quad (88)$$

$$Var\left[M_{FD} \mid H_1\right] = \frac{1}{f_s T_s}\left\{8\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^4 + \chi^4\sigma_s^4\left(2\kappa_s - 1\right)\right.$$
$$\left. + 4\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^2\left(-\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^2 + N_0\right) + 2\chi^2\sigma_s^2\left(2\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^2 + N_0\right) + N_0^2\right\}. \quad (89)$$

Proof

$$E\left[M_{FD} \mid H_1\right] = E\left[\left|y_{FD}(n)\right|^2 \mid H_1\right] = E\left[\left|\boldsymbol{w}(n) + \boldsymbol{h}_{si}(n)\boldsymbol{x}(n) + \boldsymbol{n}(n)\right|^2\right]$$
$$= E\left[|\boldsymbol{w}(n)|^2 + |\boldsymbol{h}_{si}(n)|^2|\boldsymbol{x}(n)|^2 + |\boldsymbol{n}(n)|^2 + \boldsymbol{w}(n)\boldsymbol{h}_{si}^*(n)\boldsymbol{x}^*(n)\right.$$
$$+\boldsymbol{w}(n)\boldsymbol{n}^*(n) + \boldsymbol{h}_{si}(n)\boldsymbol{x}(n)\boldsymbol{w}^*(n) + \boldsymbol{h}_{si}(n)\boldsymbol{x}(n)\boldsymbol{n}^*(n)$$
$$\left.+\boldsymbol{w}^*(n)\boldsymbol{n}(n) + \boldsymbol{h}_{si}^*(n)\boldsymbol{x}^*(n)\boldsymbol{n}(n)\right]$$
$$= 2\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^2 + \chi^2\sigma_s^2 + N_0. \quad (90)$$

On the other hand,

$$Var\left[M_{FD} \mid H_1\right] = \frac{1}{f_s T_s} Var\left[\left|y_{FD}(n)\right|^2 \mid H_1\right]$$
$$= \frac{1}{f_s T_s}\left\{E\left[\left|y_{FD}(n)\right|^4 \mid H_1\right] - \left(E\left[\left|y_{FD}(n)\right|^2 \mid H_1\right]\right)^2\right\}. \quad (91)$$

We already have $E\left[\left|y_{FD}(n)\right|^2 \mid H_1\right]$ in Eq. 90. So we only need to calculate $E\left[\left|y_{FD}(n)\right|^4 \mid H_1\right]$.

$$E\left[\left|y_{FD}(n)\right|^4 \mid H_1\right] = E\left[\left|\boldsymbol{w}(n) + \boldsymbol{h}_{si}(n)\boldsymbol{x}(n) + \boldsymbol{n}(n)\right|^4\right]$$
$$= E\left[\left||\boldsymbol{w}(n)|^2 + |\boldsymbol{h}_{si}(n)|^2|\boldsymbol{x}(n)|^2 + |\boldsymbol{n}(n)|^2 + \boldsymbol{w}(n)\boldsymbol{h}_{si}^*(n)\boldsymbol{x}^*(n)\right.\right.$$
$$+\boldsymbol{w}(n)\boldsymbol{n}^*(n) + \boldsymbol{h}_{si}(n)\boldsymbol{x}(n)\boldsymbol{w}^*(n) + \boldsymbol{h}_{si}(n)\boldsymbol{x}(n)\boldsymbol{n}^*(n)$$
$$\left.\left.+\boldsymbol{w}^*(n)\boldsymbol{n}(n) + \boldsymbol{h}_{si}^*(n)\boldsymbol{x}^*(n)\boldsymbol{n}(n)\right|^2\right]$$
$$= E\left[|\boldsymbol{w}(n)|^4 + |\boldsymbol{h}_{si}(n)|^4|\boldsymbol{x}(n)|^4 + |\boldsymbol{n}(n)|^4 + 4|\boldsymbol{w}(n)|^2|\boldsymbol{h}_{si}(n)\boldsymbol{x}(n)|^2\right.$$
$$\left.+4|\boldsymbol{w}(n)|^2|\boldsymbol{n}(n)|^2 + 4|\boldsymbol{h}_{si}(n)\boldsymbol{x}(n)|^2|\boldsymbol{n}(n)|^2\right]$$
$$= 8\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^4 + 2\chi^4\kappa_s\sigma_s^4 + 2N_0^2 + 8\left(\chi^2\sigma_s^2 + N_0\right)\sum_{i=1}^{p}\gamma_i\sigma_{w,i}^2 + 4\chi^2\sigma_s^2 N_0. \quad (92)$$

In the above expression we have used Eq. 87. By combining Eqs. 90, 91 and 92 proves Eq. 89. □

Acknowledgements This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.04-2017.16.

References

1. Mitola J (2000) Cognitive radio: an integrated agent architecture for software defined radio, Ph.D. Dissertation, KTH, Stockholm, Sweden
2. Haykin S (2005) Cognitive radio: brain-empowered wireless communications. IEEE J Sel Areas Commun 23:201–220
3. Hu W, Willkomm D, Abusubaih M, Gross J, Vlantis G, Gerla M, Wolisz A (2007) Cognitive radios for dynamic spectrum access - dynamic frequency hopping communities for efficient IEEE 802.22 operation. IEEE Commun Mag 45:80–87
4. Goldsmith A, Jafar S, Maric I, Srinivasa S (2009) Breaking spectrum gridlock with cognitive radios: an information theoretic perspective. Proc. IEEE 97:894–914
5. Wang B, Liu K (2011) Advances in cognitive radio networks: a survey. IEEE J Sel Topics in Signal Process 5:5–23
6. Hossain E, Niyato D, Han Z (2009) Dynamic spectrum access and management in cognitive radio networks. Cambridge Univ. Press
7. Akyildiz I, Lee W, Vuran M, Mohanty S (2006) Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. Comput Netw 50:2127–2159
8. Liang Y-C, Zeng Y, Peh E, Hoang AT (2008) Sensing-throughput tradeoff for cognitive radio networks. IEEE Tran Wireless Commun 7:1326–1337
9. Jafar S, Srinivasa S (2007) Capacity limits of cognitive radio with distributed and dynamic spectral activity. IEEE J Sel Areas in Commun 25:529–537
10. Tang L, Chen Y, Hines E, Alouini M-S (2011) Effect of primary user traffic on sensing-throughput tradeoff for cognitive radios. IEEE Tran Wireless Commun 10:1063–1068
11. Kim J, Shin Y, Ban TW, Schober R (2011) Effect of spectrum sensing reliability on the capacity of multiuser uplink cognitive radio systems. IEEE Trans Veh Technol 60:4349–4362
12. Zhang H, Zhang Z, Dai H (2013) On the capacity region of cognitive multiple access over white space channels. IEEE J Sel Areas in Commun 31:2517–2527
13. Sabharwal A, Schniter P, Guo D, Bliss D, Rangarajan S, Wichman R (2014) In-band full-duplex wireless: challenges and opportunities. IEEE Journal on Selected Areas in Communications 32:1637–1652
14. Korpi D, Riihonen T, Syrjälä V, Anttila L, Valkama M, Wichman R (2014) Full-duplex transceiver system calculations: analysis of ADC and linearity challenges. IEEE Trans Wireless Commun 13:3821–3836
15. Kim D, Lee H, Hong D (2015) A survey of in-band full-duplex transmission: from the perspective of phy and mac layers. IEEE Communications Surveys & Tutorials 17:2017–2046. Fourthquarter
16. Dinc T, Chakrabarti A, Krishnaswamy H (2016) A 60 ghz cmos full-duplex transceiver and link with polarization-based antenna and rf cancellation. IEEE Journal of Solid-State Circuits PP(99):1–16
17. Amjad M, Akhtar F, Rehmani MH, Reisslein M, Umer T (2017) Full-duplex communication in cognitive radio networks: a survey. IEEE Communications Surveys Tutorials 19:2158–2191. Fourthquarter
18. Li D, Cheng J, Leung VCM (2018) Adaptive spectrum sharing for half-duplex and full-duplex cognitive radios: from the energy efficiency perspective. IEEE Trans Commun 66:5067–5080

19. Li D, Zhang D, Cheng J (2019) Degrees of freedom for half-duplex and full-duplex cognitive radios. IEEE Trans Veh Technol 68:2571–2584
20. Hanawal MK, Nguyen DN, Krunz M (2020) Cognitive networks with in-band full-duplex radios: jamming attacks and countermeasures. IEEE Trans Cogn Commun Network 6(1):296–309
21. Khafagy MG, Alouini M, Aïssa S (2018) Full-duplex relay selection in cognitive underlay networks. IEEE Trans Commun 66:4431–4443
22. Cheng W, Zhang X, Zhang H (2011) Imperfect full duplex spectrum sensing in cognitive radio networks. In: Proc. of the 3rd ACM workshop on cognitive radio networks, pp 1–6
23. Cheng W, Zhang X, Zhang H (2011) Full duplex spectrum sensing in non-time-slotted cognitive radio networks. In: Proc. IEEE Military Commun. Conf. (MILCOM), pp 1029–1034
24. Afifi W, Krunz M (2015) Incorporating self-interference suppression for full-duplex operation in opportunistic spectrum access systems. IEEE Trans Wirel Commun 14:2180–2191
25. Afifi W, Krunz M (2014) Adaptive transmission-reception-sensing strategy for cognitive radios with full-duplex capabilities. In: 2014 IEEE international symposium on dynamic spectrum access networks (DYSPAN), pp 149–160
26. Liao Y, Wang T, Song L, Han Z (2014) Listen-and-talk: full-duplex cognitive radio networks. In: Global communications conference (GLOBECOM), 2014 IEEE, pp 3068–3073
27. Liao Y, Wang T, Song L, Han Z (2017) Listen-and-talk: protocol design and analysis for full-duplex cognitive radio networks. IEEE Trans Veh Technol 66:656–667
28. Wang X, Jia M, Guo Q, Ho IW, Lau FC (2019) Full-duplex relaying cognitive radio network with cooperative nonorthogonal multiple access. IEEE Syst J 13:3897–3908
29. Ali B, Mirza J, Zhang J, Zheng G, Saleem S, Wong K (2019) Full-duplex amplify-and-forward relay selection in cooperative cognitive radio networks. IEEE Trans Vehicular Technol 68:6142–6146
30. Nguyen HV, Nguyen DV, Shin OS (2020) In-band full-duplex relaying for SWIPT-enabled cognitive radio networks. Electronics 2020 9(835):1–13
31. Stotas S, Nallanathan A (2011) Enhancing the capacity of spectrum sharing cognitive radio networks. IEEE Trans Veh Technol 60:3768–3779
32. Ozcan G, Gursoy MC, Gezici S (2014) Error rate analysis of cognitive radio transmissions with imperfect channel sensing. IEEE Trans Wireless Commun 13:1642–1655
33. Ranjbar M, Nguyen-Le H, Karacolak T, Tran NH (2019) Energy efficiency of noma-based wireless networks under gaussian-mixture interference. In: ICC 2019 - 2019 IEEE international conference on communications (ICC), pp 1–6
34. Ranjbar M, Tran NH, Nguyen TV, Gursoy MC, Nguyen-Le H (2018) Capacity-achieving signals for point-to-point and multiple-access channels under non-gaussian noise and peak power constraint. IEEE Access 6:30977–30989
35. Quek TQS, de la Roche G, Güvenç I, Kountouris M (2013) Small cell networks: deployment, PHY techniques, and resource management, 1st edn. Cambridge University Press, Cambridge
36. Gulati K, Evans BL, Andrews JG, Tinsley KR (2010) Statistics of co-channel interference in a field of Poisson and Poisson-Poisson clustered interferers. IEEE Trans Signal Process 58:6207–6222
37. Das A (2000) Capacity-achieving distributions for non-Gaussian additive noise. In: Proc. IEEE Int. Symp. Inform. Theory (ISIT), p 432
38. Wiklundh K, Stenumgaard P, Tullberg H (2009) Channel capacity of Middleton's class A interference channel. Electronics Lett 45:1227–1229
39. Pighi R, Franceschini M, Ferrari G, Raheli R (2009) Fundamental performance limits of communications systems impaired by impulse noise. IEEE Trans Commun 57:171–182
40. Sanaei A, Ardakani M (2007) LDPC code design for non-uniform power-line channels, EURASIP. J Adv in Signal Process 2007:1–9
41. Ardakani M, Kschischang FR, Wei Y (2005) Low-density parity-check coding for impulse noise correction on power-line channels. In: Proc. IEEE Int. Symp. on Power Line Commun. and its Applicat., pp 90–94
42. Vu HV, Tran NH, Gursoy MC, Le-Ngoc T, Hariharan S (2015) Capacity-achieving distributions of impulsive ambient noise channels. In: Proc. IEEE Int. Conf. Commun. (ICC), pp 1–7
43. Vu HV, Tran NH, Gursoy MC, Le-Ngoc T, Hariharan S (2015) Capacity-achieving input distributions of additive quadrature gaussian mixture noise channels. IEEE Trans Commun 63:3607–3620
44. Gursoy MC (2009) On the capacity and energy efficiency of training-based transmissions over fading channels. IEEE Trans Inform Theory 55:4543–4567
45. Verdú S (2002) Spectral efficiency in the wideband regime. IEEE Trans Inf Theory 48(6):1319–1343
46. Caire G, Tuninetti D, Verdú S (2004) Suboptimality of TDMA in the low-power regime. IEEE Trans Inform Theory 50:608–620
47. Akin S, Gursoy MC (2013) On the throughput and energy efficiency of cognitive MIMO transmissions. IEEE Trans Veh Technol 62:3245–3260
48. Wang S, Wang Y, Coon JP, Doufexi A (2012) Energy-efficient spectrum sensing and access for cognitive radio networks. IEEE Trans Veh Technol 61:906–912
49. Zhang L, Xiao M, Wu G, Li S, Liang YC (2016) Energy-efficient cognitive transmission with imperfect spectrum sensing. IEEE J Sel Areas in Commun PP(99):1–1
50. Xiong C, Lu L, Li GY (2014) Energy-efficient spectrum access in cognitive radios. IEEE J Sel Areas in Commun 32:550–562
51. Mili MR, Musavian L, Hamdi KA, Marvasti F (2016) How to increase energy efficiency in cognitive radio networks. IEEE Trans Commun PP(99):1–1
52. Verdú S (1990) On channel capacity per unit cost. IEEE Trans Inform Theory 36:1019–1030
53. Yuan Y, Bahl P, Chandra R, Moscibroda T, Wu Y (2007) Allocating dynamic time-spectrum blocks in cognitive radio networks. In: Proceedings of the 8th ACM international symposium on mobile Ad Hoc networking and computing, MobiHoc '07, (New York, NY, USA). ACM, pp 130–139
54. Bany Salameh H, Krunz M (2009) Channel access protocols for multihop opportunistic networks: challenges and recent developments. Network, IEEE 23:14–19
55. Sahai A, Hoven N, Tandra R (2004) Some fundamental limits in cognitive radio. In: Allerton Conf. Commun. Controls and Comput.
56. Tandra R, Sahai A (2005) Fundamental limits on detection in low SNR under noise certainty. In: Proc. IEEE Int. Conf. on Wireless Networks, Commun. and Mobile Comput., pp 464–469
57. Tandra R, Sahai A (2008) SNR Walls for signal detection. IEEE J Sel Areas in Commun 2:4–17
58. Wang T, Liao Y, Zhang B, Song L (2015) Joint spectrum access and power allocation in full-duplex cognitive cellular networks. In: 2015 IEEE International Conference on Communications (ICC), pp 3329–3334
59. Boulogeorgos AA, Salameh HAB, Karagiannidis GK (2017) Spectrum sensing in full-duplex cognitive radio networks under hardware imperfections. IEEE Trans Veh Technol 66:2072–2084
60. Li W, Lilleberg J, Rikkinen K (2014) On rate region analysis of half- and full-duplex OFDM communication links. IEEE Journal on Selected Areas in Communications 32:1688–1698

61. Li L, Dong C, Wang L, Hanzo L (2016) Spectral-efficient bidirectional decode-and-forward relaying for full-duplex communication. IEEE Trans Veh Technol 65:7010–7020
62. Gaafar M, Amin O, Abediseid W, Alouini M (2017) Underlay spectrum sharing techniques with in-band full-duplex systems using improper Gaussian signaling. IEEE Trans Wirel Commun 16:235–249
63. Pinsker MS, Prelov VV, Verdu S (1995) Sensitivity of channel capacity. IEEE Transactions on Information Theory 41(6 pt 2):1877–1888

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

M. Ranjbar[1] · H. L. Nguyen[2] · N. H. Tran[1] · T. Karacolak[3] · S. Sastry[1] · L. D. Nguyen[4]

M. Ranjbar
mr130@zips.uakron.edu

N. H. Tran
nghi.tran@uakron.edu

T. Karacolak
tutku.karacolak@wsu.edu

S. Sastry
ssastry@uakron.edu

L. D. Nguyen
dinhlonghcmut@gmail.com

1 Department of Electrical and Computer Engineering, University of Akron, Akron, OH, USA

2 The University of Danang, University of Science and Technology, Danang, Vietnam

3 School of Engineering and Computer Science, Washington State University Vancouver, Vancouver, WA, USA

4 Duy Tan University, Danang, Vietnam