RESEARCH

On the feasibility of a secondary service transmission over an existent satellite infrastructure: design and analysis

Luciano Barros Cardoso da Silva[1,2]* ![ORCID], Tarik Benaddi[1] and Laurent Franck[3]

Abstract

In this paper, we present a realistic use case in order to investigate the feasibility of a secondary service transmission over an existent satellite infrastructure. By introducing the overlay cognitive radio paradigm towards satellite communications, we compute a theoretical achievable data rate greater than 16 kbps for the secondary service, which is suitable for most M2M applications. Using simulation results, we show that this can be achieved while preserving the primary service performance. In addition, a system design framework is discussed in order to dimension such systems.

Keywords: Satellite communications, Cognitive radio, Overlay paradigm, Dirty paper coding, Machine-to-machine application

1 Introduction

It can be emphatically stated that the access to space is easier, cheaper, and faster than ever before. This is the widespread view among the main satellite players with respect to the unique opportunistic time currently experienced by the space segment. Actually, the favorable projections to this segment could be sustained especially today, since the demand for the rising new services has increased considerably [1] (multicast, broadcast, high mobility, and wide coverage).

In this sense, as an actual and typical example, it could be pointed out the use of satellite to support the machine-to-machine (M2M) communications, providing connectivity to the end-users anytime, anywhere, for any media and device [2]. The M2M communications are one of the central use cases in the fifth-generation (5G) mobile network [3] as they play a major role in the Internet of Things (IoT). In fact, it is predicted in [4] the deployment of around 1 million devices per kilometer square like sensors/actuators, vehicles, factory machines, and beacons.

In what concerns the technological advances, this satellite era consolidation could be reinforced (but not only) by the recent maturity reached in the manufacturing process (cheaper and faster production, powerful and sophisticated payloads). To keep up with the new challenges of M2M communications, the satellite communication systems need to push the boundaries in the direction to more efficient technical solutions. For this purpose, the search for power and bandwidth efficiency as well as the actual trend to low complexity systems must be a paramount concern for system designers.

For a given provided service to the end-users, the design requirements of the communication system, in general, are settled by (i) the service availability, which ultimately specifies the required bit error rate (BER); (ii) the allocated spectral band, which is assigned by the International Telecommunication Union Radio Regulation [5] (ITU-RR), and (iii) the range of the received carrier-to-noise ratio (CNR), measured in the occupied bandwidth, which is limited by the large path loss as well as the nonlinear behavior of the satellite channel. For instance, it could be typically assigned the dynamic range of CNR between

*Correspondence: luciano.barroscardosodasilva@imt-atlantique.fr

[1]IMT Atlantique, LabSTICC, DEOS, 10 Avenue Edouard Belin, BP 44004, 31028 TOULOUSE CEDEX 4 Toulouse, France

[2]National Institute for Space Research, INPE, Av dos Astronautas, 1758 -. Jardim da Granja, 12227-010 São José dos Campos, Sao Paulo, Brazil

Full list of author information is available at the end of the article

— 3 and 20 dB, in accordance to the extension recently presented at the DVB-S2X standard [6].

Basically, concerning these bounds, the system can be designed efficiently in (i) power, by decreasing the received power (or equivalently the CNR) necessary to reach the specified BER, for instance by adding redundancy bits in a digital encoding system, or (ii) spectral efficient, by increasing the number of bits per hertz into the occupied bandwidth, which ultimately increases the transmitted data rate. Last but not least, the (iii) system complexity should be carefully evaluated in order to reduce as much as possible the number of processing operation, especially when on-board systems are considered.

Apart from that, since the available radio spectrum is today a scarce resource (cf [7] for example), another challenge faced is to develop techniques which enable a better coordination between legacy and future services, especially considering this new machine-type communication environment and its large-scale implementation. Despite of the spectrum regulation and policy still being considered as a dry subject since the earliest days of radio communication, the need to reconsider the static long-term exclusivity of the spectrum via licensed regulation procedures, as well as the encouragement of the techniques which enable the coexistence of different networks, become a key element to permit a proper expansion of these new services [8].

It is within this framework that the cognitive radio (CR) techniques have also become attractive for space applications [9]. Based on the recent developments in the space qualified software-defined radios (SDR) [10] and also by the acceptability of concepts such as flexible [11, 12] and hosted payloads, these techniques allow a smarter spectrum management. In addition, some valuable research [13] has been acted in spectral awareness and spectral exploitation techniques, which has driven the cognitive satellite communications towards a promising approach.

In a nutshell, the cognitive user (CU), in our context, unlicensed or less prioritized to operate in a specific spectrum band, senses the environment around it, and adapts its transmission as a function of the interference, by adjusting the frequencies, waveforms, and protocols in order to access the licensed primary user (PU) spectrum efficiently. Without going into further details, three paradigms classify the CU operation [14]:

- *Interweave*, which is based on the idea of opportunistic transmission. In short, the CU observes the *white* spaces not used by the PU transmission, (in space, time or frequency) and adjusts its operating parameters. It is inferred that, ideally, there is no coexistence between users and, consequently, no power control of the CU transmission is required;

- *Underlay*, denominated as *gray* spectrum space, where the CU, by means of the partial knowledge of the PU signal characteristics and channel, adjusts its parameters in such a way as to respect an acceptable interference threshold. As a greater knowledge of the PU signal is required, more sophisticated spectral sensing techniques must be employed, such as signal-to-noise ratio (SNR) estimation. It is worth noting that the *interweave* paradigm can be seen as a special case of *underlay*, where the threshold does not allow any interference. Some examples of techniques found in the literature are: dynamic resource control (power and frequency), beamforming with multiple antennas, and spectral spreading;

- *Overlay*, where the CU, from the full and noncausal knowledge about the PU waveform and message and channel, uses advanced coding and modulation strategies to transmit simultaneously while mitigating the interference. The occupied spectrum space in this last paradigm is called *black*, due to the fact that it is occupied by the interfering signals and noise.

The first two schemes were well studied in [13, 15]. In this paper, we investigate the third scheme.

The main reason to propose the overlay paradigm for satellite communications lies in the feasibility of transmitting both unlicensed and licensed services simultaneously from the same satellite towards its respective terminals. We emphasize that, due to the priority among users, the superposition coding strategy is required [16] to relay the PU transmission, unlike the technical solutions adopted for the broadcast channel [17]. Furthermore, the dirty paper coding (DPC) [18] is implemented to adapts the cognitive signal to the direction of the PU interference.

Concerning the DPC, from a brief historical perspective, the first idea of practical scheme was proposed by Erez, Shamai and Zamir [19]. It pointed out the Tomlinson-Harashima precoding (THP) for intersymbol interference (ISI) canceling, which can be seen as a DPC application for frequency selective channel. In this technique, the modulo operation is used to pre-subtracted the interference with a minimum power increase. Also in this work, the THP losses, i.e., shaping loss at high signal-to-noise ratio (SNR) regime and the combined modulo and power losses at intermediate/low SNR regimes, were well characterized. Moreover, Eyuboglu and Forney in their seminal paper [20] generalized the combination of the trellis shaping (TS) [21], trellis-coded modulation (TCM), and THP for Gaussian ISI channels. The so-called trellis precoding (TP) performs interference pre-subtraction and allows recovery of the shaping loss. Likewise, a little bit more closer of our application, an extension of TP for multiuser interference was proposed to recover the shaping loss with sufficient high constellation expansion in [22] and [23],

where the TS technique acts as a vectorial quantization, replacing the modulo operator.

In view of the above, the techniques previously exposed in the recent publications [24, 25] concerns the design of the overlay paradigm transmission towards satellite communication systems.

This paper is an extended version of the author's work [26]. Additionally, to the link budget evaluation and system dimensioning introduced in [26], we provide here a design framework to implement such cognitive systems as well as a detailed analysis of the effect of different parameters on the overall system performance. Firstly, a DPC encoder is proposed involving TS and TCM concepts along with proper constellation expansion combined with THP. The discussions leads to a trade-off between power efficiency, by the reduction of the modulo loss, and complexity, key levers for both satellite onboard processing and terminals. Secondly, as a general contribution, we focus on the feasibility of a low data rate secondary transmission. In this sense, a practical use case is investigated, which considers commercial off-the-shelf (COTS) parts [27] and assumes realistic link budget parameters in its evaluation. The discussions and results contained herein could be seen as part of a "preliminary phase" of an engineering process plan [28].

Following this introduction, Section 2 presents the overlay model description. Subsequently, in Section 3, the employed methodology is introduced, with reference to the superposition and dirty paper coding (DPC) designs. The paper contributions on these techniques are detailed in Section 4, emphasizing the satellite context. Next, the results and discussion are presented, by investigation of a realistic use case in Section 5. Finally, Section 6 is dedicated to the conclusions, with suggestions for future works.

2 Overlay model description

The following scenarios are provided as examples where the overlay CR techniques might be applied to satellite communications. In the first case, presented in the Fig. 1a, an ordinary low Earth orbit (LEO), or medium Earth orbit (MEO), satellite provides two different services towards different terminals. In this context, a single licensed user PU takes priority over the added unlicensed CU. The interference presented at both terminals should be mitigated by properly designed CU encoder, without any changes in the PU transmission chain.

In the same way, the GEO multibeam satellite is illustrated in the Fig. 1b. In this case, considering the frequency reuse, the CU is able to transmits by using, for instance, the determined blue frequency (or polarization) into the red spot footprint, as far as the interference among adjacent beams is resolved. It is worth noting also that all possible different PU transmissions, represented by several blue spots, should be taken into account in the interference mitigation design. By this way, the total satellite capacity could be increased as well as the spectrum resources better managed.

Equally suitable for both scenarios, the interference model with side information, adapted from [16], is presented in Fig. 2. Assuming that the signals are onboard the satellite, the cognitive encoder has full and noncausal knowledge about each PU ith signal and message, which addresses the main overlay paradigm requirement. In this sense, the encoded cognitive signal X_c^n is function of both primary and cognitive messages $m_{p,i}$ and m_c.

Without loss of generality, considering the ith PU and the added CU, the channel gains $| h_{yx,i} |$ (from the transmitter x to the receiver y) are defined by the direct paths ($| h_{cc} |$ and $| h_{pp,i} |$), and the interfering paths ($| h_{pc,i} |$ and $| h_{cp,i} |$) losses (the Fig. 2 summarizes these notation). In our context, these gains are computed as function of the each transmission link budget.

The following equations describes the output of the channel, where n refers to the nth symbol:

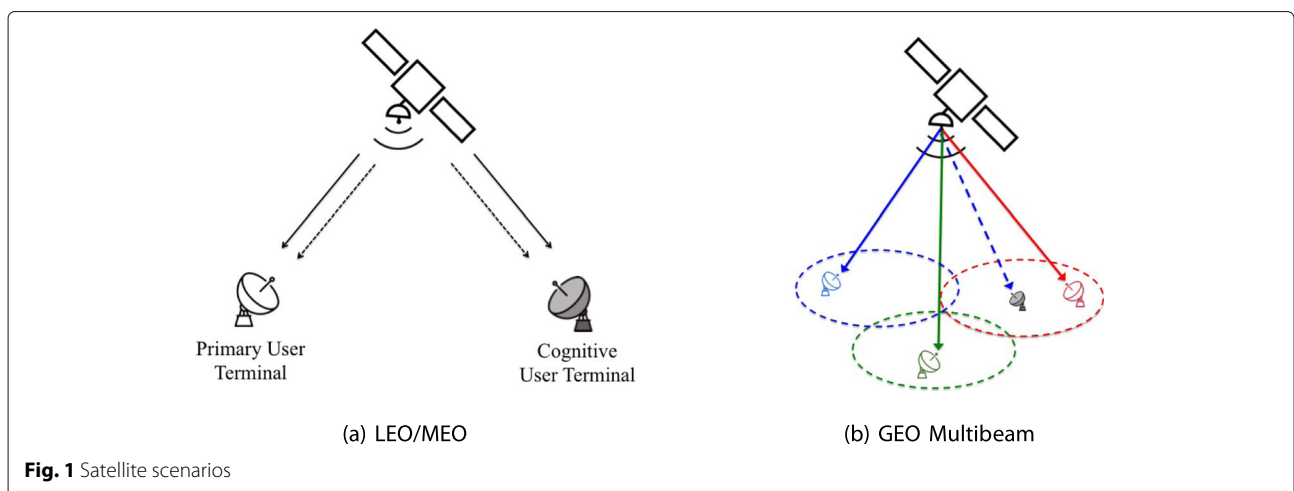$$Y_{p,i}^n = | h_{pp,i} | X_{p,i}^n + | h_{pc,i} | X_c^n + Z_{p,i}^n \qquad (1)$$
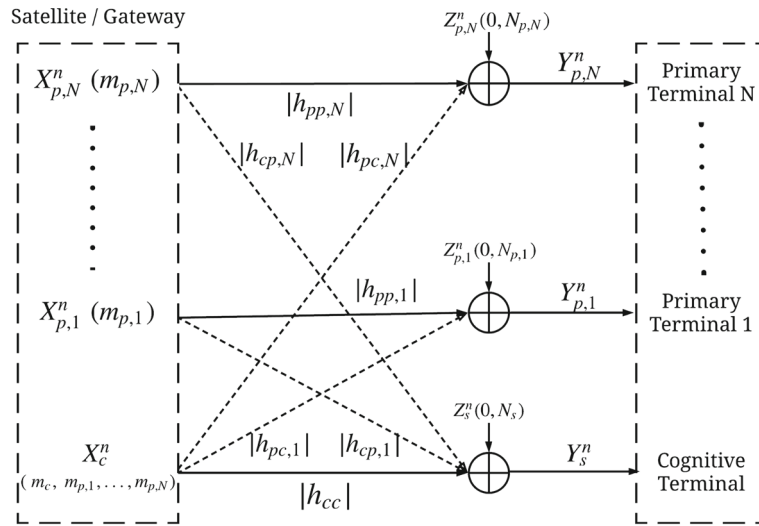

Fig. 1 Satellite scenarios

(a) LEO/MEO   (b) GEO Multibeam

Primary User Terminal

Cognitive User Terminal

Fig. 2 Overlay model

$$Y_s^n = \mid h_{cc} \mid X_c^n + \sum_{i=1}^{N} \mid h_{cp,i} \mid X_{p,i}^n + Z_s^n. \tag{2}$$

Based on the fact that the terminals may be located in different geographical sites, the Gaussian noise component $Z_{p,i}^n$ (resp. Z_s^n) is assumed to follow the normal law $\mathcal{N}\left(0, N_{p,i}\right)$ (resp. $\mathcal{N}(0, N_s)$). Also, the power constraints to be satisfied are $E\left[\|X_{p,i}^n\|^2\right] = P_{p,i}$ and $E\left[\|X_c^n\|^2\right] \leq P_c$, respectively.

Finally, since each PU has the same transmission priority, we highlight that the interference among them could be solved by precoding techniques as proposed, for instance, in DVB-S2X standard [6]. Under this assumption, this work provides a design method to permit a secondary service transmission without affecting the PU transmission performance.

3 Methodology
3.1 Superposition strategy
The purpose of the superposition technique is to ensure that the signal-to-noise ratio (SNR) at each PU receiver is not decreased in the presence of interference. To accomplish this goal, the CU shares part of its power to relay each PU message, and $\hat{X}_c^n$ is a modified version of the CU message, as it is going to be detailed in the next section. Based on that operation, the CU transmitted signal is given by

$$X_c^n = \hat{X}_c^n + \sum_{i=1}^{N} \sqrt{\alpha_i \frac{P_c}{P_{p,i}}} X_{p,i}^n, \tag{3}$$

where $\alpha_i \in [0, 1]$ is the shared power fraction from P_c to relay each PU message.

Under the assumption that all signals are statistically independent, the new power constraint can be defined as $E[\|\hat{X}_c^n\|^2] \leq \left(1 - \sum_{i=1}^{N} \alpha_i\right) P_c$. The signal-to-interference-plus-noise ratio (SINR) at the kth primary receiver is given by

$$\text{SINR}_{p,k} = \frac{E\left[\left\|\left(\mid h_{pp,k} \mid + \mid h_{pc,k} \mid \sqrt{\alpha_k \frac{P_c}{P_{p,k}}}\right) X_{p,k}^n\right\|^2\right]}{E\left[\left\|\mid h_{pc,k} \mid \left(\hat{X}_c^n + \sum_{i=1,i\neq k}^{N-1} \sqrt{\alpha_i \frac{P_c}{P_{p,i}}} X_{p,i}^n\right)\right\|^2\right] + E\left[\left\|Z_{p,k}^n\right\|^2\right]}$$

$$= \frac{\left(\mid h_{pp,k} \mid \sqrt{P_{p,k}} + \mid h_{pc,k} \mid \sqrt{\alpha_k P_c}\right)^2}{\|h_{pc,k}\|^2(1 - \alpha_k)P_c + N_{p,k}} \tag{4}$$

In this context, the superposition factor $\alpha_k \in [0, 1]$ that maximizes Eq. (4), for the interference condition ($\mid h_{pc,k} \mid > 0$), which is a generalized form of [16, Eq. 14], is given by

$$\alpha_k = \left(\frac{\mid h_{pp,k} \mid \sqrt{P_{p,k}}\left(\sqrt{N_{p,k}^2 + \|h_{pc,k}\|^2 P_c(N_{p,k} + \|h_{pp,k}\|^2 P_{p,k})} - N_{p,k}\right)}{\mid h_{pc,k} \mid \sqrt{P_c}(N_{p,k} + \|h_{pp,k}\|^2 P_{p,k})}\right)^2 \tag{5}$$

Hence, by applying Eq. (5) in Eq. (4), we have

$$\text{SINR}_{P,k} = \frac{\|h_{pp,k}\|^2 P_{p,k}}{N_{p,k}} \tag{6}$$

By inspection of Eq. (3), we emphasize that the CU transmission is feasible only if the condition $\sum_{i=1}^{N} \alpha_i < 1$ is satisfied. By this assumption, note that the CU data rate should be decreased when aggressive frequency reuse scenarios are considered.

3.2 Dirty paper coding

Once the superposition factors are computed and the CU partially shares its power to relay each PU signal, the next step is to design $\hat{X}_c^n$ efficiently, in order to minimize the interference of PU transmission on the CU receiver. The optimal strategy employes the theoretical results presented by Costa [18]. On the assumption that the interference (PU signal) is noncausally known at the CU transmitter, a transmitter-based interference presubtraction can be implemented, without any power increase, reaching the AWGN capacity.

By rearranging the Eq. (2) and considering the superposition, we have

$$Y_s^n = |h_{cc}| \hat{X}_c^n + \sum_{i=1}^{N} \left(|h_{cp,i}| + |h_{cc}| \sqrt{\alpha_i \frac{P_c}{P_{p,i}}} \right) X_{p,i}^n + Z_s^n. \tag{7}$$

Without loss of generality, given that the signals in Eq. (7) are statistically and mutually independents, the implemented model considers a single Gaussian distributed PU constellation, in respect to the total interfering power received at CU terminal. In addition, in order to simplify the notation through this paper, the Eq. (7) is normalized by the direct path attenuation factor $|h_{cc}|$. Thus, the signal at CU receiver is given by

$$Y_s^n = \hat{X}_c^n + \underbrace{\left(b + \sqrt{\alpha \frac{P_c}{P_p}} \right) X_p^n + Z_s^n}_{S^n}, \tag{8}$$

where the factor b represents the normalized interfering path and S^n represents the total channel interference.

Figure 3 presents the basic diagram of the DPC encoder, where the THP is used to presubtract the multiuser interference. In this configuration, assuming low and intermediate SNR regime, the partial interference presubtraction (PIP) is implemented [18]. In this way, the signal $\hat{X}_c^n$ is designed as

$$\hat{X}_c^n = \left[X_{cc}^n - \lambda S^n \right] \text{MOD}_\Delta , \tag{9}$$

where X_{cc}^n is the coded signal and the factor λ, to be properly chosen, controls the fractioned interference to be presubtracted. Also, MOD_Δ is the complex-valued modulo operation, implemented to limit the transmitted power. The modulo region is defined by $\Delta = \sqrt{M}d_{\min}$, where M is the number of points of the square QAM constellation and $d_{\min}$ is the minimum intersymbol distance.

When compared to QAM signal on AWGN channel, the performance of THP system is degraded by three separated sources of losses: *power loss*, *modulo loss*, and *shaping loss* [22]. From this point, our goal is to design the coded signal X_{cc}^n in order to mitigate these sources of losses.

3.3 Trellis shaping-based DPC
3.3.1 Main encoding blocks

Figure 4 presents the implemented encoder for the cognitive user. Three gains can be achieved by this system: (i) coding gain, represented by the upper part of the diagram; (ii) shaping gain, achieved by the trellis shaping code in the lower part, and (iii) prediction gain (term rubricated by [29]), achieved by the modulo operation jointly with shaping code.

The two codes work independently. The input bit sequence is split in two parts. The upper part is formed by the coset select code C_c. This later is described by the generator matrix G_c, which encodes the k_c message bits x_c into a n_c-coded bits y_c. In the lower part, the r_s-bits syndrome sequence s passes through an inverse syndrome former H_s^{-T} for the shaping code C_s. This initial sequence t jointly with the channel coded sequence y_c and the interference λS^n are fed into the *Viterbi* decoder. This later selects, according to a well chosen branch metric, the shaping codeword y_s. After that, the shaped sequence z_s is obtained by the XOR operation between t and y_s. Note that z_s and t are within the same coset, according to the trellis shaping on regions strategy, detailed in [21]. Finally, the output shaped sequence X_{cc}^n is obtained by mapping the d symbols as a function of y_c and the sign mapped bits z.

Similar to the THP operation, the coded symbol sequence X_{cc}^n is presubtracted by the scaled interference sequence λS^n and modulo operated in order to limit
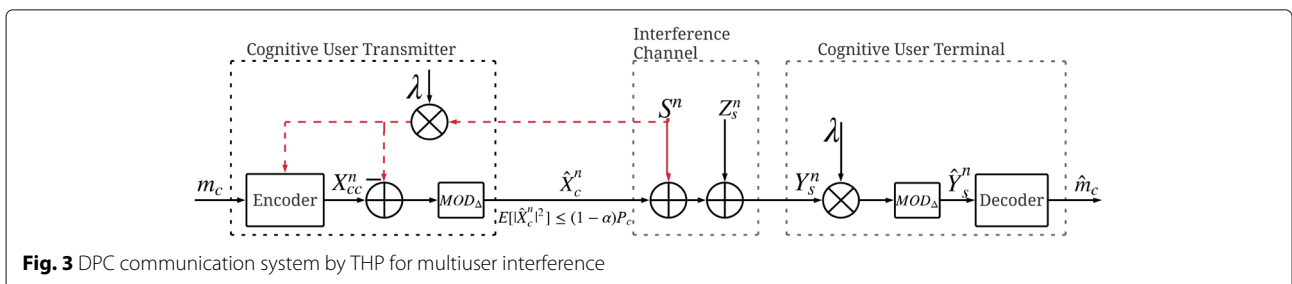

Fig. 3 DPC communication system by THP for multiuser interference
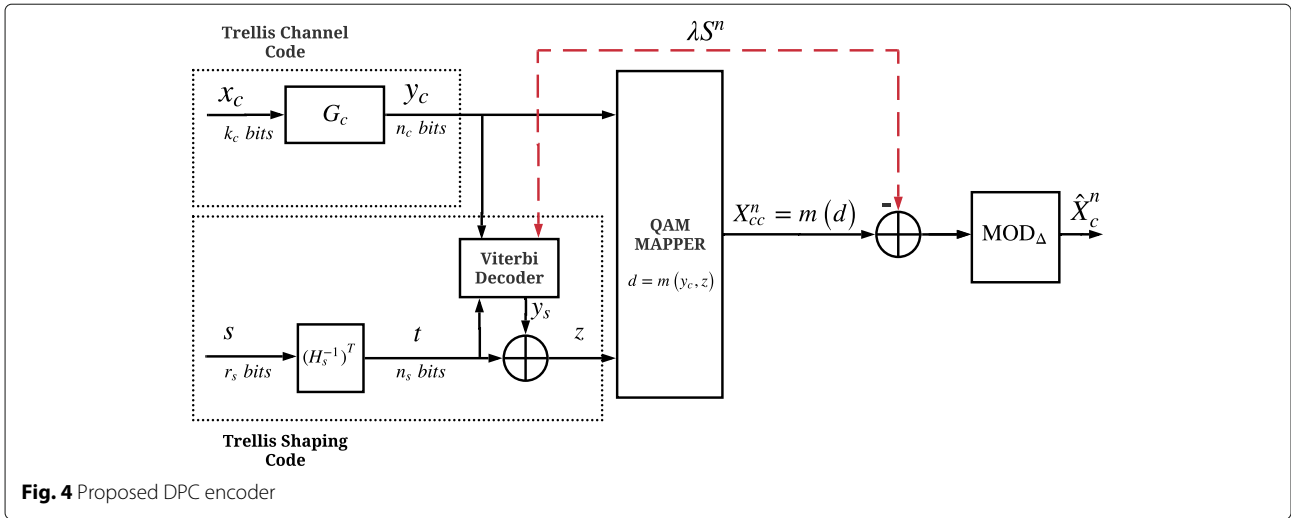
Fig. 4 Proposed DPC encoder

its power. At the end, the transmitted sequence $\hat{X}_c^n$ is obtained and sent through the interference channel.

The following branch metric Eq. (10) is implemented, where the precoder selects the proper region sequence with minimum average power to steer the scaled interference sequence λS^n:

$$\|[X_{cc}^n - \lambda S^n]\text{MOD}_\Delta\|^2. \tag{10}$$

3.3.2 Transmitter design

We propose, in this section, a scheme based on the trellis shaped DPC encoder with slight constellation expansion combined with THP. Unlike the previous works published in [22, 23], the modulo operation is always part of our transmission system, assuring the required power limitation for the transmitter, given any interference scenario.

The available remaining transmitted power (after the superposition) is computed based on the link budget parameters. In order also to comply with the defined SNR range for satellite communication, it is proposed a typical transmission rate of $R_{cu} = 2$ bits/symbol. The trellis shaped-based DPC is implemented, with a slight expanded constellation of $n_s = 2$, and modulo operation in its output, previously depicted in Fig. 4.

For the coding gain, a systematic 64-state, rate 1/2, convolutional code C_c specified in octal notation by the feedforward polynomial $h_1(D) = 54$ and the feedback polynomial $h_0(D) = 161$ is assumed. Along with this, for the shaping code C_s, the 4-state code, rate 1/2, specified by generators $g_{s,1}(D) = 7$ and $g_{s,2}(D) = 5$ is implemented. In other words, at the DPC transmitter, the 4-QAM-coded constellation is replicated four times by the shaping operation (since $n_s = 2$), resulting in an expanded 16-QAM constellation for X_{cc}^n.

In addition, considering the shaping operation, the proposed mapping for X_{cc}^n is based on the sign bit shaping strategy described by Forney in [21]. Following the same

notation as in Fig. 4, consider now the 16-QAM X_{cc}^n constellation represented in Fig. 5. Each symbol is defined by the tuple $(z_1 z_2 y_{c_1} y_{c_2})$, where z_1 and z_2 are the sign bits while y_{c_1} and y_{c_2} are the coset bits. It should also be noted that, to keep the specified rate, which maintains the SNR range, the uncoded bits are not implemented in this scheme (when compared with the general scheme presented in the previously [22]).

Thus, the shaping regions are represented by the four different colors that delimit the constellation regions, while the four different markers $\bigcirc$, $\times$, $\square$, or $\triangle$ represent the convolutional code cosets. The analogy with the *Gelfand - Pinsker* multicoding scheme [30], which is a key concept utilized in the DPC theoretical proof [18], is
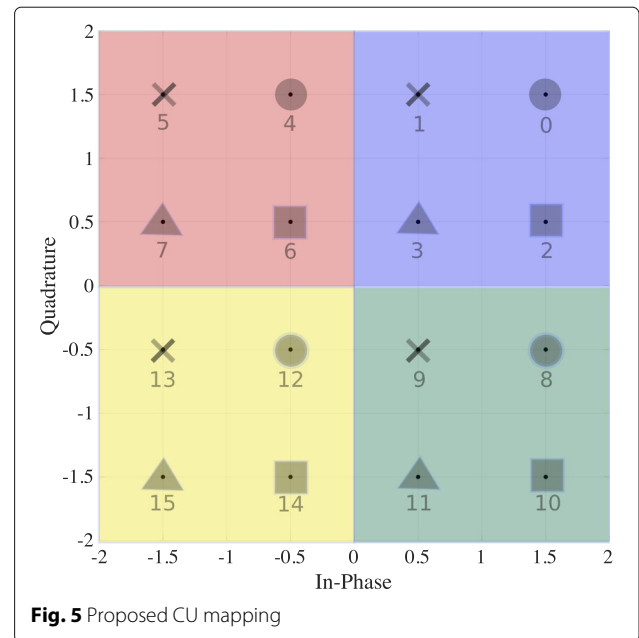

Fig. 5 Proposed CU mapping

intuitive. Basically, we assign the transmitted "sequence of colors" as the subcodebook, determined by the shaping code C_s, which is indexed by the chosen codeword generated by the convolutional code C_c. In this sense, the described encoder can be conceptually considered as a practical implementation of the DPC encoder.

3.3.3 Receiver design

At the receiver, as depicted in the Fig. 6, the reverse chain is implemented: firstly, Y_s^n is multiplied by the factor λ. Before entering the DPC decoder, the signal is modulo operated again, as the following:

$$
\hat{Y}_s^n = \left[(Y_s^n)\lambda \right]_{\mathrm{MOD}_\Delta}
$$

$$
= \left[\lambda \hat{X}_c^n + (\hat{X}_c^n - \hat{X}_c^n) + \lambda S^n + \lambda Z^n \right]_{\mathrm{MOD}_\Delta} \quad (11)
$$

$$
= \left[\lambda \hat{X}_c^n + [X_{cc}^n - \lambda S^n]_{\mathrm{MOD}_\Delta} - \hat{X}_c^n + \lambda S^n + \lambda Z^n \right]_{\mathrm{MOD}_\Delta}
$$

$$
= \left[X_{cc}^n - (1 - \lambda)\hat{X}_c^n + \lambda Z_s^n \right]_{\mathrm{MOD}_\Delta}, \quad (12)
$$

where in Eq. (11), the following property was utilized $[(a)\mathrm{MOD}_\Delta + b]\,\mathrm{MOD}_\Delta = (a + b)\mathrm{MOD}_\Delta$.

The value of λ that minimizes the effective noise $N_{eff} = (1 - \lambda)\hat{X}_c^n + \lambda Z_s^n$ is obtained by [18]:

$$
\lambda = \frac{(1 - \alpha)P_c}{(1 - \alpha)P_c + E\left[|Z_s^n|^2 \right]}. \quad (13)
$$

The decoder employs the same strategy for usual TCM schemes. Figure 7 illustrates an example of trellis for the channel code C_c with the parallel transitions. Since each branch of the trellis corresponds to a signal subset (in this case, defined by y_c and labeled by the markers $\bigcirc$, X, $\square$, or $\triangle$), the first step in decoding is to determine the best signal point within each subset (equivalently to determine the best shaping region, represented by the different colors). This is performed by selecting the point that is closest in Euclidean distance to the received point $\hat{Y}_s$. After that, the selected signal point and its squared distance metric are applied in the usual *Viterbi* decoder for C_c in order to select the most likely coded sequence $\hat{y}_c$ [31]. At the end of the decoding, the estimated shaping bits $\hat{s}$ are obtained by $\hat{s} = \hat{z}H_s^T$.

By observing this simple example, the trellis transition from the state σ_j^n to the state σ_j^{n+1} provides $(y_{c_1}y_{c_2})$ mapped to $(0, 0)$, which is represented by the marker "$\bigcirc$" according with the assigned mapping. Thus, the decoder will perform a hard decision among the four constellation symbols $(0, 4, 8, 12)$, representing the shaping regions, defined by the pair of bits $(z_1 z_2)$. In the same way, the transition from the state σ_j^n to σ_{j+1}^{n+1} provides $(y_{c_1}y_{c_2})$ mapped to $(0, 1)$, represented by the marker "X," which decides among the parallel transitions 1, 5, 9, or 13.

It is worth noting that an error in the first step of the decoding procedure (i.e., error in the parallel transition decision) results in a wrong decoding of the shaping region. This fact increases significantly the degradation caused by the modulo loss in DPC schemes, as will be further discussed in the next section. We also remark that, in order to reduces the system complexity, the state expansion in the *Viterbi* decoder at the receiver [32–34], which could be an effective strategy for eliminating the parallel transition effect, is not considered in this implementation.

4 CU practical encoder design for dirty paper channel

In this section, we describe our contributions based on some changes in the low complexity encoder presented in the last section. Firstly, taking into account that the CU signal is Gaussian distributed by trellis shaping operation, we demonstrate a procedure that provides the appropriate output power for $\hat{X}_c^n$. As a consequence, the SINR of the PU is properly evaluated. Subsequently, we deal with the CU link by proposing further expansion in the X_{cc} constellation, jointly with an optimized mapping design. The results are analyzed as a function of the complexity involved in the transmitter implementation.

4.1 CU transmitted power control

The power reduction of $\hat{X}_c^n$, caused by the trellis precoding technique (i.e., $E[|\hat{X}_c^n|^2] < (1 - \alpha)P_c$), impacts directly on both link performances since the value of α, adopted for power sharing at the superposition strategy, is no longer exact. It can be noted, as given by Eq. (6), that
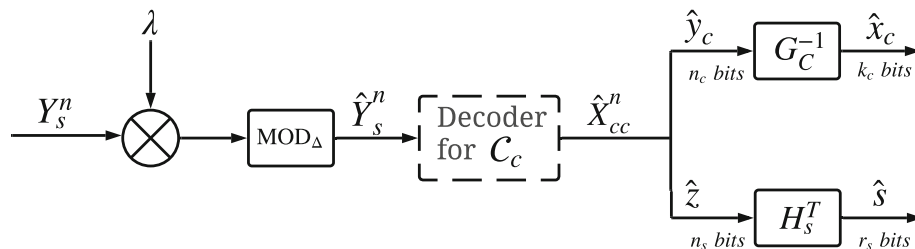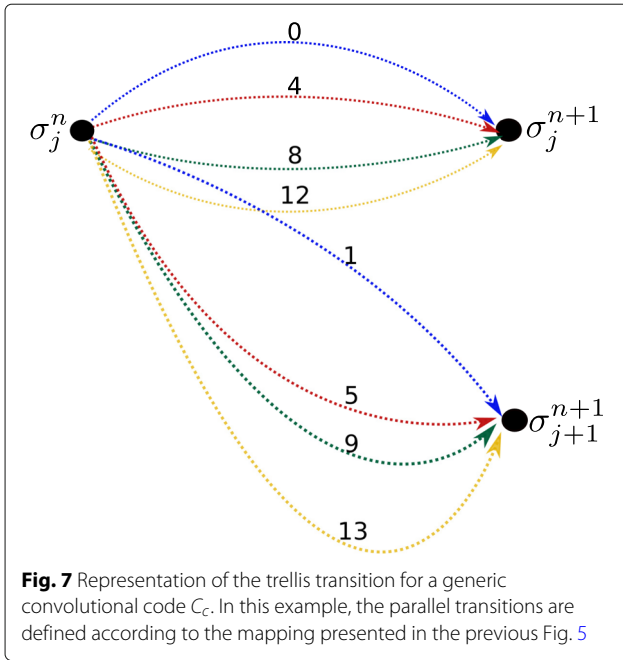

Fig. 6 Proposed DPC receiver

Fig. 7 Representation of the trellis transition for a generic convolutional code C_c. In this example, the parallel transitions are defined according to the mapping presented in the previous Fig. 5

the SINR at primary receiver is increased. As a result, the PU presents better bit error rate (BER) performance than when operating in AWGN channel (as further analyzed in [25]).

By considering that the exactly same shaping gain generated by a shaping code C_s is obtained for the multiuser precoding [22], we propose a method for controlling the CU output power such as $E[|X_c^n|^2] = P_c$, or equivalently, $E\left[|\hat{X}_c^n|^2\right] = (1-\alpha)P_c$. This is reached by proper scaling the minimum distance d_{min} of the coded constellation X_{cc}^n.

As a reference, we assume the power of the baseline constellation, without considering the shaping operation, given by [35]:

$$P_\oplus = \frac{2^R}{6}d_{min}'^2, \tag{14}$$

where R is the data rate in bits per two dimensions, without taking into account the shaping redundant bits. The scaled minimum distance is defined as d_{min}'. The shaping gain is then defined as

$$\gamma_s = \frac{P_\oplus}{E\left[|\hat{X}_c^n|^2\right]} = \frac{P_\oplus}{(1-\alpha)P_c}. \tag{15}$$

Thus, by combining Eq. (14) and Eq. (15) and rearranging the terms, the d_{min}' such that the available power after the shaping operation is equals to $(1-\alpha)P_c$, is given by

$$d_{min}' = \sqrt{\frac{[(1-\alpha)P_c]\,6\gamma_s}{2^R}}. \tag{16}$$

All things considered, some remarks are pointed out for this proposed procedure:

1. *Both links are properly adjusted.* Since the output power of the remaining CU transmission, after superposition, satisfies exactly $E\left[|\hat{X}_c^n|^2\right] = (1-\alpha)P_c$, the BER performance of PU in a interference channel is the same as AWGN channel. In addition, as the CU power is efficiently employed, the BER of cognitive link is improved or, equivalently, the secondary service data rate can be augmented;

2. *CU modulo loss might be reduced.* The performance degradation caused by modulo loss augments as a function of: (i) the selection probability of boundaries symbols in the X_{cc}^n constellation and (ii) the reduction of the SNR level. By these perceptions, it is worth noting that, by re-scaling the minimum distance of X_{cc}^n and considering the same Gaussian distributed interference S, this procedure might reduce the occurrence of boundaries selected points for X_{cc}^n, since the constellation region is further enlarged. Clearly, because of this modulo loss reduction, this technique improves the cognitive system performance for the same fixed SNR (in the next section, the investigation of a practical use case will further demonstrate this concept).

3. *Mapping of X_{cc}^n constellation should be optimized.* In order to obtain exactly the specified output power $(1-\alpha)P_c$, the shaping operation should be performed over a defined continuous region delimited by the modulo amplitude Δ. In this sense, it is of utmost importance that the mapping be designed such that the interference presubtracted signal $\hat{X}_c^n$ be confined within this delimited region. Otherwise, the previous procedure does not control the exact power. This specific issue, jointly considered with constellation expansion, is further discussed in the next section.

4.2 Constellation expansion

The THP precoding works as the simplest solution for multiuser interference (MUI) presubtraction. By using the modulo operation, this technique satisfies the power constraint for application of dirty paper encoder and results in almost negligible degradation in high SNR regimes [36–38].

However, for low and intermediate SNR regimes (for instance, below 15 dB), the degradation caused by THP losses becomes more significant, especially due to the modulo loss. Additionally, the modulo loss increases as a function of the probability occurrence of the boundary constellation points in X_{cc}^n.

In order to recover part of the THP losses, some practical precoding techniques are proposed in [22]. For intermediate SNR regimes, which is our application, [22] proposes the combination of TP and PIP techniques. The

recommended shaping code is 5/6 in order to expand the X_{cc}^n, in such way to confine the interference. Also, as the trellis shaping is implemented at the precoder, the transmitted presubtracted signal $\hat{X}_c$ has Gaussian distribution. It is interesting to observe that, as the modulo operation is not considered at the precoding, only the power and shaping losses affect the system performance.

As a complement of the approach presented in [22], we consider the modulo implemented at the transmitter output in our investigated schemes. The amount of the interference is function of the link budget and, on the contrary of [22], different constellation expansions are considered. This assumption allows the analysis of the modulo loss impact. Also, the results could be evaluated considering the trade-off between complexity and power efficiency.

As an example for our proposed constellation expansion, and following the same notations in Fig. 4, let us consider that the original DPC constellation is a 16-QAM. Each symbol is defined by the tupple $(z_1 z_2 y_{c_1} y_{c_2})$, where z_1 and z_2 are the sign bits while y_{c_1} and y_{c_2} are the coset bits (as proposed in our low complexity design in the Fig. 5). This constellation can be expanded to, for instance, a 64-QAM by considering an additionally two information bits u_1 and u_2 or two *"auxiliary"* bits (not information bits) z_{aux_1} and z_{aux_2}. Now, each symbol is defined by the tupple $(z_1 z_2 u_1 u_2 y_{c_1} y_{c_2})$ or $(z_1 z_2 z_{aux_1} z_{aux_2} y_{c_1} y_{c_2})$, respectively.

By employing the same procedure, the signal constellation of X_{cc}^n can be expanded as necessary to confine the scaled interference (i.e., $|\operatorname{Re}(\lambda S^n)|$ and $|\operatorname{Im}(\lambda S^n)| < (\sqrt{M}/2).d_{\min}$, where M is the constellation order of X_{cc}^n). In this way, the assignment of the boundary constellation points can be avoided and, as consequence, the modulo loss is mitigated. Also, it is highlighted that the original information rate is maintained by expanding with auxiliary bits, since these last ones are not information

bits. For the transmitter, this operation could be seen as an extension of the trellis shaping procedure. The *Viterbi* decoder of the shaping operation acts as an usual TCM decoder, where the auxiliary shaping bits z_{aux} represent the parallel branches transitions. These latter are hard decided during each trellis section for C_s.

Furthermore, it is still necessary to design the optimal mapping of shaping regions, according to the information rate. The next section discusses this topic.

4.3 Mapping discussion

Consider the implementation of the mapping I, depicted in the Fig. 8. We assume a shaping code C_s of rate 1/2, which provides four shaping regions, represented by the different colors. The convolutional code C_c of rate 1/2 is defined, which also provides four TCM cosets. In addition, the modulo regions are represented by dashed lines, according to the X_{cc} constellation expansion adopted in 16, 64, 256, and 1024-QAM.

The interference point is represented by the red "X", which is presubtracted by the X_{cc} constellation point, selected by C_s, to form the transmitted $\hat{X}_c$ symbol. As we have discussed in the Section 4.1, by a proper mapping design, we confine the $\hat{X}_c$ signal in a determined continuous region of the constellation. In this way, the control of the transmitted power is achieved. Following, two cases are analyzed.

In the first case, depicted in the Fig. 8a, the uncoded information bits in the trellis precoding are considered. We define the mapping, according to the sign bit shaping strategy, proposed in [21], as $(z_1 z_2 u_1 u_2 u_3 u_4 u_5 u_6 y_{c_1} y_{c_2})$. In this way, we have 8 information bits transmitted by symbol, which require high SNR (unlike the operating range targeted in this article). Thus, there is no shaping subregions in this scheme. Each 256-QAM constellation, divided in 4 cosets with 64 constellation points inside
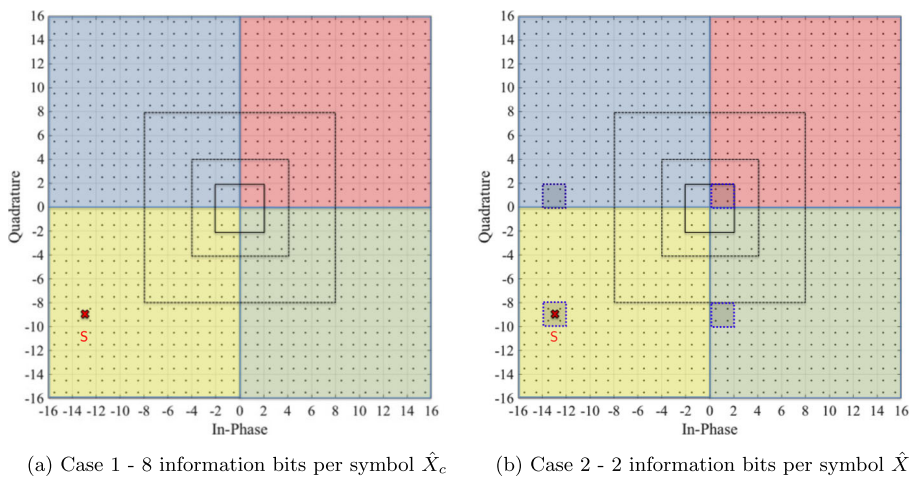

(a) Case 1 - 8 information bits per symbol $\hat{X}_c$ (b) Case 2 - 2 information bits per symbol $\hat{X}_c$

Fig. 8 Shaping region mapping I

each coset, is repeated 4 times to form the outer shaping 4-QAM. Regardless of whether the interference S is confined within the 1024-QAM constellation of X_{cc}, the presubtracted signal $\hat{X}_c$, due to the modulo, is always restrict in its region. Based on this assumption, we observe the following:

- *The shaping loss is partially recovered.* This is reached according to the shaping gain γ_s, provided by the code C_s, and based on the fact that $\hat{X}_c$ is continuous and shaped bounded by the squared region of $\Delta = 32$ (i.e., $|\operatorname{Re}(\hat{X}_c)|$ and $|\operatorname{Im}(\hat{X}_c)| < \Delta/2$, where Δ is the modulo boundary);
- *The modulo loss is assumed negligible.* In fact, even if there are some points selected at the boundaries of X_{cc}, since the system operates in high SNR regime, the modulo loss only cause degradation in very low BER;
- *The power loss is close to 0.* By the continuous approximation [35], we compute the power loss as $\gamma_p = 10.\log(1024/1023) \approx 0.0042$ dB for this 1024-QAM constellation.

In the second case, depicted in Fig. 8b, we transmit our rate of 2 information bits per $\hat{X}_c$ symbol. The uncoded bits are replaced by the auxiliary shaping bits and the mapping is defined by $(z_{aux_1} z_{aux_2} z_{aux_3} z_{aux_4} z_{aux_5} z_{aux_6} (z_{aux_1} z_{aux_2} z_{aux_3} z_{aux_4} z_{aux_5} z_{aux_6}$ $z_1 z_2 y_{c_1} y_{c_2})$. The shaping operation selects the closest one of the 2^6 subregions (or equivalent, points inside each shaping cosets) of each one of the 4 region defined by C_s.

Considering this rate, as discussed in our proposed power control technique, the $\hat{X}_c$ signal must be continuous, Gaussian shaped, and bounded by the squared region of $\Delta = 4$ (i.e., $|\operatorname{Re}(\hat{X}_c)|$ and $|\operatorname{Im}(\hat{X}_c)| < \Delta/2$, where Δ is the modulo boundary). For the same interference S, represented also by the red "X," the only case that the presubtracted signal $\hat{X}_c$ is confined in our region of interest occurs when the "yellow" region is selected by C_s. Intuitively, as the other shaping regions (i.e., colors) can be selected by C_s, it is not guaranteed that the shaping gain provided by this code is reached. As a result, the power control strategy, discussed in the last section, would not be effective in this case.

The following mapping, represented in Fig. 9, is optimized for scenarios where auxiliary bits are employed. In this case, the mapping is defined by $(z_{aux_1} z_{aux_2} z_{aux_3} z_{aux_4} z_{aux_5} z_{aux_6} (z_{aux_1} z_{aux_2} z_{aux_3} z_{aux_4} z_{aux_5} z_{aux_6}$ $z_1 z_2 y_{c_1} y_{c_2})$. The same four shaping cosets are represented by the four different colors and the 2^6 subregions (points inside each shaping coset) are spread over the expanded constellation.

The interference point is also represented by the red "X", which is presubtracted by the closest X_{cc} to provide the transmitted signal $\hat{X}_c$. It is worth noting that, considering
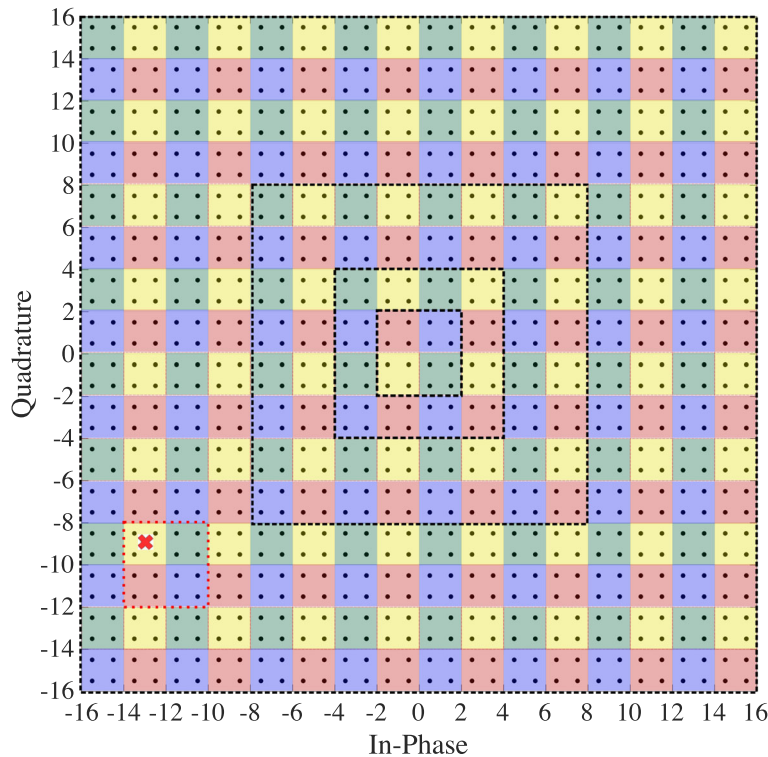

Fig. 9 Shaping regions mapping II

the hard decision of the parallel branches for the C_s trellis, $\hat{X}_c$ is always inside the smallest dashed square, which is our continuous region of interest ($\Delta = 4$). By this design, we conclude that

- *The shaping loss is partially recovered.* the shaping gain γ_s is reached according to the code C_s. As a consequence, the power control technique is properly designed;
- *The modulo loss can be totally mitigated.* If the X_{cc} constellation is expanded enough such that the interference is confined and the boundary selected symbols are avoided, the modulo loss can be mitigated. However, we emphasizes that, in some cases, a low modulo loss might be tolerated in order reduces the transmitter complexity (this will be discussed in the next section, by analyzing the realistic scenario);
- *The power loss.* Considering this rate, this loss is evaluated by the approximation $\gamma_p = 16/15 \approx 0.28$ dB. As discussed in [22], this loss represents the relation between continuous, which is the DPC case, and discrete constellation transmission;

In summary, the previous discussions showed, by considering proper expansion of the DPC constellation, that the degradation can be made within 0.3 dB of the corresponding reference trellis shaping in AWGN channel. On other hand, we have seen that the system complexity increased, which is a design drawback. It is important to emphasize that the PU user maintains the same performance as in absence of the CU operation.

In this work, we investigate the constellation expansion in 4, 16, 64, and 256 regions, which represents $n_s = 2, 4, 6, 8$ (where, in case of $n_s > 2$, the remaining $n_s - 2$ are auxiliary bits). It is important to point out that the minimum distance $d'_{\min}$, evaluated by scaling method, previously exposed in Section 4.1, was implemented in all simulated schemes. Additionally, the impact of modulo operation in the system is also considered.

As we are employing $R_{cu} = 2$ bits/symbol, the transmitted signal should be confined within the region defined by the 16-QAM constellation (i.e., $|\operatorname{Re}(\hat{X}_c)|$ and $|\operatorname{Im}(\hat{X}_c)| < 2d_{\min}'$). Having this in mind, we are assuming the previously discussed mapping II (see Fig. 9) for all considered schemes.

The following example, depicted in Fig. 10, clarifies the mapping strategy. In this case, the constellation is expanded from 16 to 64 regions, where the X_{cc} symbol is defined by $(z_{\text{aux1}}, z_{\text{aux2}}, z_1, z_2, y_{c_1}, y_{c_2})$. Consequently, we have the middle bits z_1 and z_2 as the shaping coset bits, which assign the shaping regions (represented by colors) and z_{aux1} and z_{aux2} as a points within each shaping
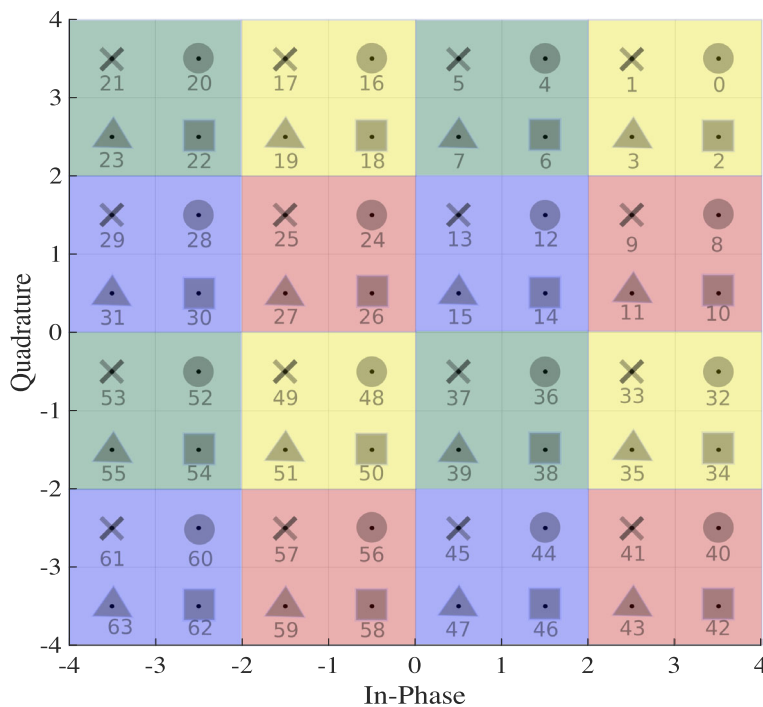

Fig. 10 Shaping regions mapping II for 64-QAM

coset, which defines the shaping subregions. Finally, the convolutional coded bits y_{c_1} and y_{c_2} selects a point (represented by the markers $\bigcirc$, X, $\square$, and $\triangle$) inside each subregion.

Consider that the pair of bits (y_{c_1}, y_{c_2}) are mapped to $(0, 0)$ by the channel trellis convolutional code C_c, represented in the figure by the marker $\bigcirc$. In this case, Fig. 11 elucidates the *Viterbi* decoder of the shaping operation located at the transmitter. By observing this example, the trellis transitions from the state σ_j^n to the state σ_j^{n+1} corresponds to constellation points where $z_1 = 0$ and $z_2 = 0$ (yellow regions). Thus, the decoder will perform a hard decision among the four constellation symbols $(0, 16, 32, 48)$, representing the points within the assigned shaping region coset, each defined by the pair of bits (z_{aux1}, z_{aux2}). In the same way, the transition from the state σ_j^n to σ_{j+1}^{n+1} corresponds to constellation points where $z_1 = 0$ and $z_2 = 1$ (green regions), which decides among the parallel transitions $(4, 20, 36, 52)$.

We can noticed that, thanks to this mapping strategy, regardless of the interference S^n, the presubtracted signal $\hat{X}_c^n$ is confined within the 16-QAM region, assuring that the power control is established. In our analyzed scenarios, when required to mitigate the modulo loss, further expansions are implemented following the same presented procedure.

4.4 Practical system analysis

This section intends to point out some practical issues concerning the techniques exposed and applied in the satellite scenario.

As a matter of system engineering, the design for the CU payload could either be a standalone system (implemented by a dedicated transmission chain and antenna) or a shared transmitter (by using the same transponder and antenna as the PU). In this latter configuration, notice that more caution should be taken into account when the transmission of both signals inputs the same high power amplifier (HPA). In fact, this practice should be avoided since this implementation may induce higher non-linear distortions, particularly in terms of AM/AM and AM/PM conversions [39].

Moreover, at the receiver side (PU and CU), two scenarios could be encountered: (i) by the deployment of geographically separated receiving sites for each user, and thus reducing the interference due to the attenuation at the interfering paths, or (ii) by using the same receiving station with two dedicated demodulators and decoders. In this last case, the attenuation of the interfering and direct paths are the same (i.e., $| h | = | h_{pp,i} | = | h_{cc} | = | h_{pc,i} | = | h_{cp,i} |$). In fact, it increases the interference of both links and, as a consequence, requires higher value of the superposition factor α (consequently, reducing the secondary service data rate).
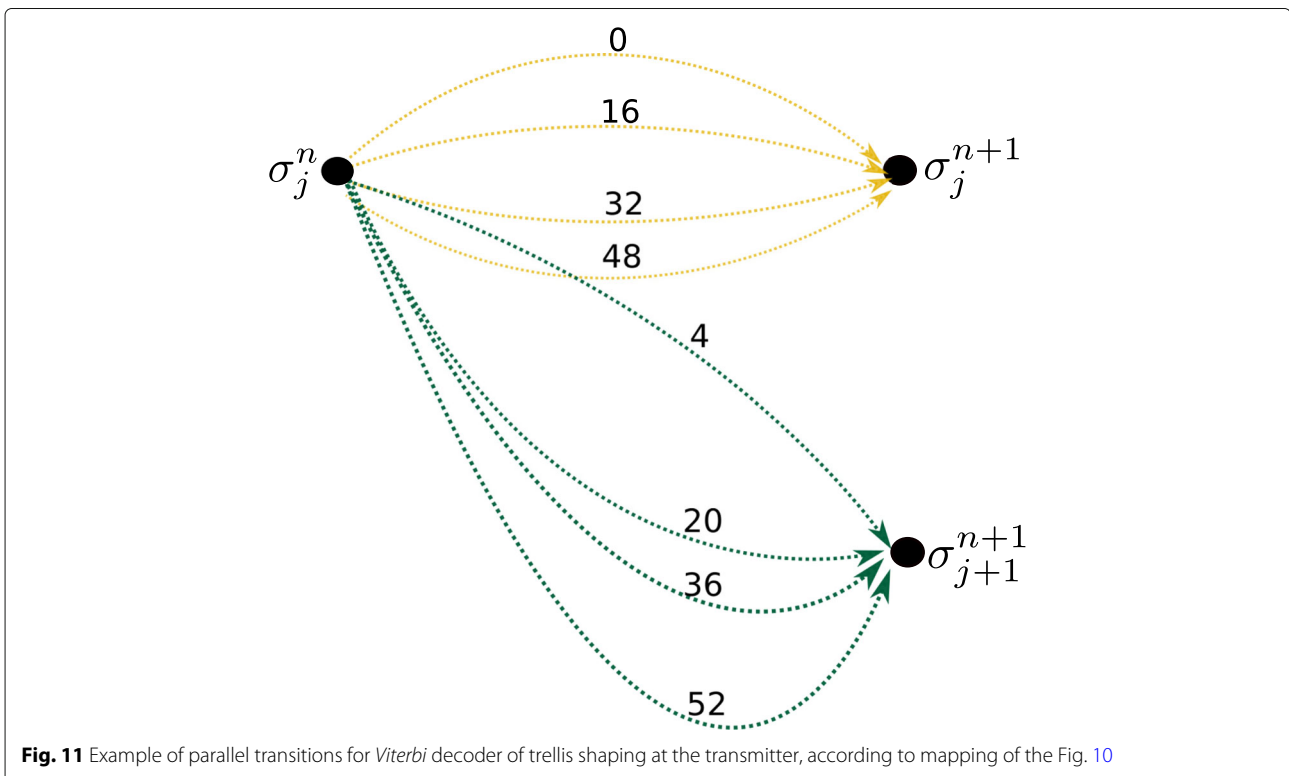

Fig. 11 Example of parallel transitions for *Viterbi* decoder of trellis shaping at the transmitter, according to mapping of the Fig. 10

Deepening the vision on the techniques described, we point out that, due to the superposition, the bit rate of the secondary service might be very low with respect to the primary. However, this practice generates two implementation problems: (i) in the DPC presubtraction technique, the same symbol rate for both signals is considered in order to be able to compute the Eq. (9) and (ii) in the superposition technique, the interference generated by the CU signal would appear as spikes in the PU bandwidth, which makes the usual interference model unrealistic in this case. In order to avoid both constraints, we can think of the implementation of the chirp spread spectrum technique [31] at CU transmission. In this sense, the DPC encoder can correctly perform its operation and the CU receiver can demodulate at a more flexible transmitted data rate.

To improve the whole system performance, the channel estimation techniques could be realized at the terminal's end through a link feedback, for instance, according to the DVB-S2X standard [6]. By these features, the superposition factor α, which depends directly on the channels conditions, as well as the λ, which depends on SNR, can be periodically updated, changing the achievable secondary service data rate and, as consequence, optimizing CU performance.

5 Results and discussion

In order to investigate the system feasibility, we adopted a scenario where a Cubesat at a altitude of 600 km with same orbital parameters as [40], using COTS parts, transmits both signals (primary and cognitive) from the same satellite antenna towards a single earth station, which is equipped with two dedicated demodulators. In this sense, the channel attenuations are the same and defined as $|h|$. In this study, just the downlink is considered.

5.1 Primary service transmission analysis

The main specification for the PU signal are output power of $1W$ [27], operating frequency of 2200 MHz (downlink band assigned for Earth exploration satellite service), bit rate of 3.4 Mbps, BER specified to 10^{-5} and coded QPSK modulation with FEC ($R = 1/2$). Table 1 presents the link budget of PU without secondary service addition.

It is worth noting that a conservative margin for demodulation losses of 6 dB is assumed in order to cover the impairments of the communication chain. The overall link margin is about 3.5 dB, as required by the targeted BER.

The principle behind this design strategy was to use part of the power remaining in this margin to transmit the CU signal. Therefore, we defined that 900 mW were allocated for PU transmission (which still maintains the recommended link margin of 3 dB) and 100 mW were used for CU. Figure 12 presents the overlay model considering this use case.

Table 1 Primary user link budget (QPSK coded FEC $R = 1/2$)

Frequency (MHz)	2200
Throughput rate (Mbps)	3.4
Transmitted power (mW)	1000
Satellite carrier EIRP (dBm)	38.3
Free space loss (dB)	−162.2
Depointing loss (dB)	−10
E. S. antenna max gain - 5m, eff 50% (dBi)	38.2
System Noise temperature (K)	130
C/N0 (dB–Hz)	81.8
Eb/N0 (dB)	16.5
Demodulation losses (dB)	− 6
Eb/N0 required (dB) – for BER = 1E-5	7
Margin (dB)	3.5

The powers transmitted and received are provided considering an example of a realistic link budget parameters. By computations according the Table 1, the channel attenuations equal $\|h\|_{\text{dB}}^2 = -125.7$ dB. In this condition, the interference-to-noise ratio (INR) and the link degradation D by interference are given by

$$\text{INR} = \frac{I_{pc}}{N_p} = 4.47; \tag{17}$$

$$D(\text{dB}) = 10.\log(1 + \text{INR}) = 7.38 \text{ dB}. \tag{18}$$

Considering all parameters, the CU performs the superposition strategy and, by the Eq. (5), the factor $\alpha = 0.85$ is evaluated. This value guarantees a SINR of 16 dB at PU. We highlight that, thanks to the superposition strategy and the power controlling design of the DPC encoder (see results in [25]), the PU maintains the same performance as in the absence of the CU interference.

5.2 Secondary service transmission analysis

From this point, a simulation for CU link is performed considering the whole CU channel interference (see Eq. (8)), which is composed by the PU signal and the CU shared power in the superposition.

5.2.1 CU results and analysis

The system performance is analyzed for the scenario depicted in Fig. 12 in both qualitative and quantitative approaches. The main objective is to provide a comparative study regarding the different expansions of X_{cc}^n, emphasizing the trade-off between power efficiency vs. complexity. Figure 13 presents the main results.

1. Qualitative analysis: To support the analysis, the scatter plot of the signal constellations, that are part of the partial interference pre-subtraction (PIP) processing are interpreted (represented in Fig. 13a,
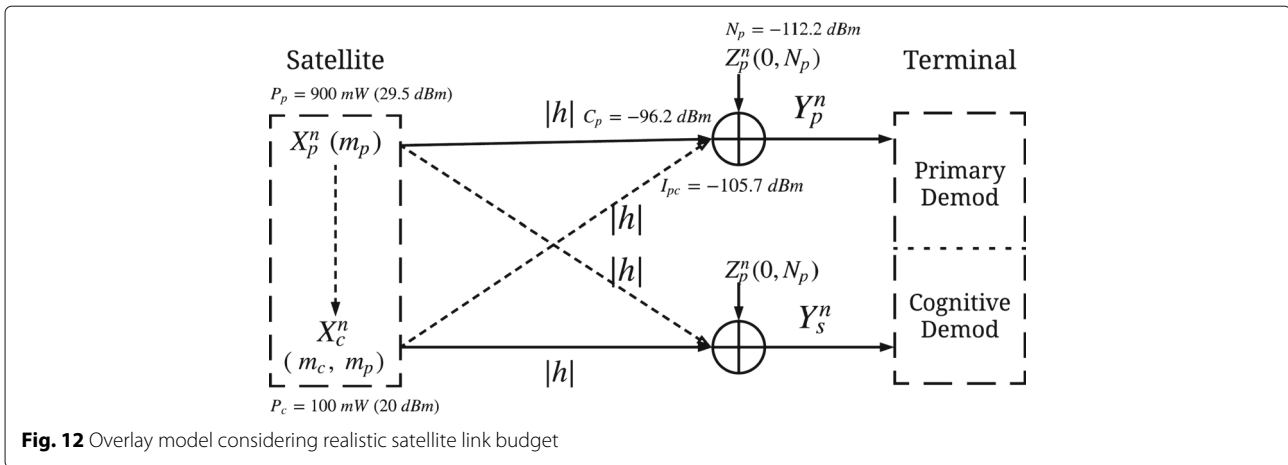
Fig. 12 Overlay model considering realistic satellite link budget

b). The expanded constellation signal X_{cc}^n is shown in green "*x*." The Gaussian distributed version of the scaled interference λS^n is superposed in red points and the transmitted signal $\hat{X}_c^n$ is shown in blue dots. Additionally, we depict the histogram of X_{cc}^n (resp. Fig. 13c, d) in order to study the probability of the boundaries symbols, since that it is determinant for the degradation caused by modulo loss. As our previous definition and in contrast to [22] and [23], we are using the modulo operation at the transmitter output and the shaping metric is implemented taking that into account Eq. (10). This design assumptions assures that the transmitted signal $\hat{X}_c^n$ will be confined within the expanded constellation, independently of the channel interference power, guaranteeing the condition $E[|\hat{X}_c^n|^2] \leq (1-\alpha)P_c$, which is a DPC requirement. By this design, in a clearly reference to the THP scheme, three important consequences are highlighted concerning its discriminated losses:

- Shaping loss: When the shaping metric takes into account the modulo operation (see Eq. (10)), the shaping gain corresponding to C_s is also reached within the continuous region delimited by Δ. It is worth noting that the same shaping gain is achieved, independently of the channel interference power. This is directly given by the fact that

$$\left[X_{cc}^n - \lambda S^n\right]\mathrm{MOD}_\Delta = \left[X_{cc}^n - (\lambda S^n)\mathrm{MOD}_\Delta\right]\mathrm{MOD}_\Delta,$$
(19)

where we can conclude that the shaping minimization will be the same whether or not λS^n is confined within the modulo region. This could be observed by the blue dots illustrated in the scatter plot, which provides the shaping gain equals to 0.97 dB for all scenarios, according to the implemented shaping code C_s. Another

interesting point concerning the shaping operation refers to the fact that, when the transmitted signal $\hat{X}_c^n$ is Gaussian distributed, thanks to the shaping operation, the usual decoding method by minimum euclidean distance is optimum [22]. It could be inferred directly due to the Gaussian distribution of the effective noise N_{eff}, given by Eq. (11).

- Modulo loss: Based on the previous observation that the boundary points should be avoided. We observe in Fig. 13c that for 16-QAM the X_{cc}^n constellation is uniformly distributed (i.e., the so-called *n-cube* distribution). We also notice that the histogram of 1024-QAM, presented in Fig. 13d, shows the occurrence of boundary constellation points decreasing as further expansion is considered.
- Power loss: The same degradation value evaluated for 16-QAM (in this case, equals to 0.28 dB), as discussed in Section 4.3.

2. Quantitative analysis: Connected with this qualitative investigation, the BER CU is shown for quantitative performance evaluation. Figure13e presents the BER curves for this scenario. We observe that the modulo loss is significantly high for the 16-QAM, 64-QAM, and 256-QAM schemes. However, when 1024-QAM is considered, the degradation just becomes more significant after 5 dB of SNR. This could be inferred by the increase of λ as a function of SNR (see Eq. (13)), which results in more modulo loss degradation. When comparing minimum and maximum X_{cc}^n constellation expansion (resp. 16-QAM and 1024-QAM), we detected a gain around 4 dB considering a BER of 10^{-3} and 1.5 dB considering a BER of 10^{-5}. This defines the minimum (by 16-QAM scheme implementation) and maximum (by 1024-QAM scheme implementation) supported

(a) Scatterplot of X_{cc}^n as 16-QAM

(b) Scatterplot of X_{cc}^n as 1024-QAM

(c) Histogram of X_{cc}^n as 16-QAM

(d) Histogram of X_{cc}^n as 1024-QAM
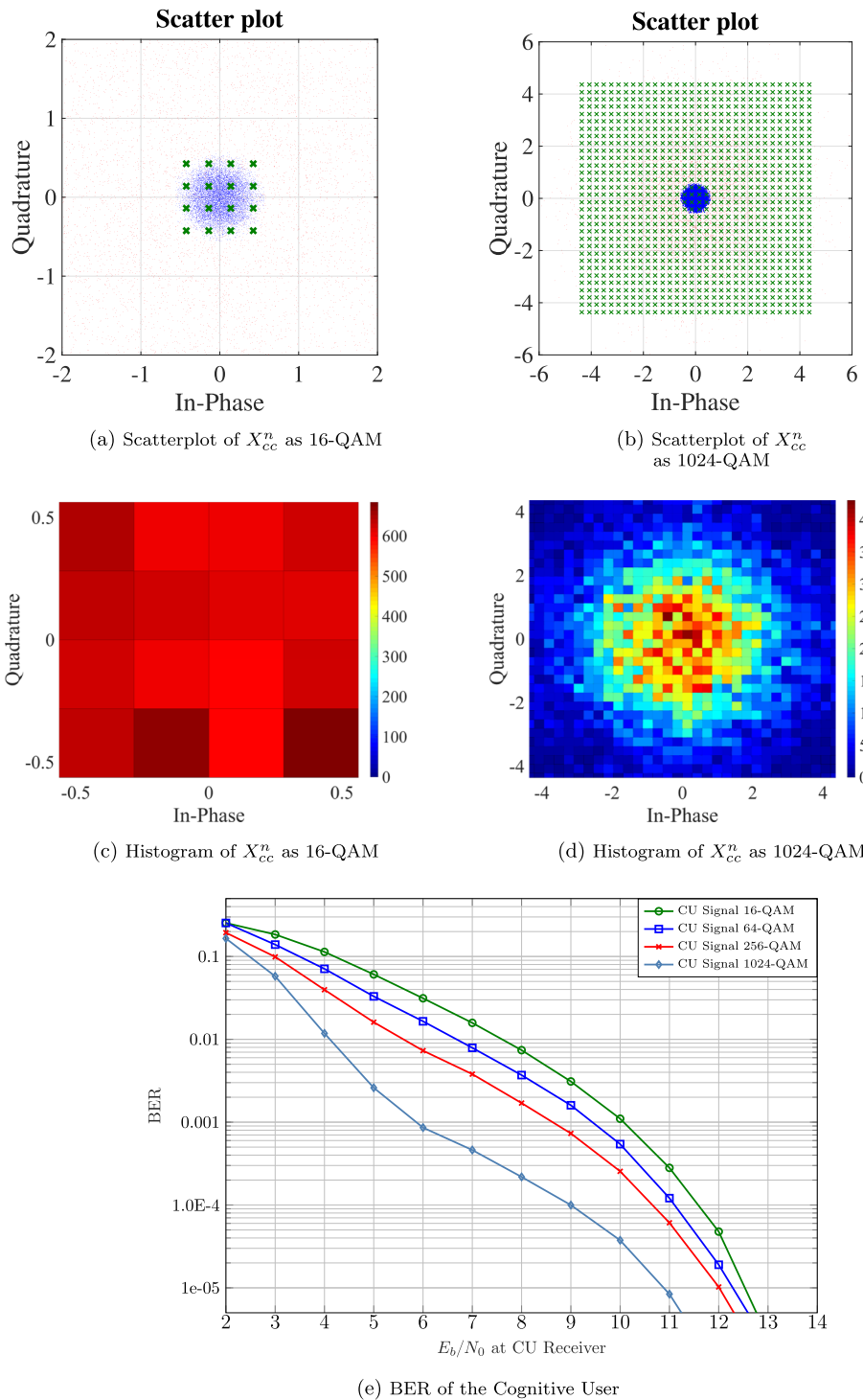
(e) BER of the Cognitive User

Fig. 13 Results and analysis for CU, considering $10.\log(Pp/Pc) = 9.5$ dB and $\alpha = 0.85$, according to different constellation expansions

bit rate for the secondary service. As a result, we point out the tradeoff between power efficiency (or alternatively the supported bit rate) and complexity. The implementation of 1024-QAM at the transmitter side could be seen as prohibitive in reference to

satellite communications. However, we emphasize that the impact is only on the increase of the system complexity, since more comparisons should be performed by the shaping operation. As we have discussed in Section 4.3, the $\hat{X}_c^n$ signal is continuous

within the squared region bounded by $\Delta = 4d_{\min}$ (i.e., $|\operatorname{Re}(\hat{X}_c)|$ and $|\operatorname{Im}(\hat{X}_c)| < \Delta/2$, where Δ is the modulo boundary). Thus, the transmitted power (i.e., $E[|\hat{X}_c^n|^2]$) as well as the power peak to average power ratio (PAPR) remain in the same order, regardless the constellation expansion.

Finally, the maintenance of the PU service is evaluated by attesting the exact specified SINR_P (see Eq. (4)) at the primary receiver, thanks to the superposition strategy.

5.2.2 CU link budget

By taking the link parameters and the CU BER curve presented in Fig. 13e, we can now evaluate the link budget, presented at Table 2. It is important to note that, by reason of superposition, just 15% of the power originally allocated for CU is used for its own transmission. However, despite of the received low power, the sensitivity of the receiver is in line with the specifications usually attributed for small satellites links (i.e., sensivity threshold of -118 dBm [41]). We emphasize that all conservative margins are still being considered in order to guarantee the performance.

It is worth noting that different M-QAM schemes present the peak to average power ratio (PAPR) in the same order of magnitude (as presented in Table 2). Also, the further expansion impacts directly in the BER performance, even when the total interference is not confined inside the expanded constellation (as in this case). As pointed in [24], by practical effects, the SNR here does not consider the defined effective noise in Section 11. As an important result of this feasibility analysis, we assure the minimum supported bit rate of 16 and 28 kbps (*w.r.t.* BER of 10^{-5} and 10^{-3}, respectively) and the maximum of 22 and 74 kbps (*w.r.t.* BER of 10^{-5} and 10^{-3}, respectively) for the secondary service. This range is suitable for most of M2M applications.

6 Conclusion

The motivation behind this paper concerns the general trend to reconsider the regulation procedure for spectrum allocation, as well as to stimulate the development of techniques which enable the coexistence of different networks. In this view, the terrestrial cognitive radio techniques have also been employed in the satellite communication context.

A step forward is presented here, where the coexistence among primaries and cognitive users is investigated by using *overlay* paradigm. In this specific paradigm, the cognitive user (CU) has noncausal knowledge about the message and encoding strategy of the primary user (PU). By this assumption, the optimum CU encoding strategy consists in: the superposition and dirty paper coding (DPC) techniques.

Table 2 Cognitive user link budget

Transmitted power (mW)	100*0.15
Satellite carrier EIRP (dBm)	20.1
Free space loss (dB)	−162.2
Depointing loss (dB)	−10
Ground station antenna max gain - 5m, eff 50% (dBi)	38.2
System noise temperature (K)	130
C (dBm)	−113.9
C/N0 (dB–Hz)	63.6
Demodulation losses (dB)	−6
Eb/N0 required (dB) – for BER = 1E-3	
DPC 16-QAM	10
DPC 64-QAM	9.4
DPC 256-QAM	8.5
DPC 1024-QAM	5.9
Eb/N0 required (dB) – for BER = 1E-5	
DPC 16-QAM	12.5
DPC 64-QAM	12.25
DPC 256-QAM	12
DPC 1024-QAM	11
Margin (dB)	3
Minimum Bit Rate (kbps) – DPC 16-QAM	
BER = 1E-3	28.8
BER = 1E-5	16.2
Maximum bit rate (kbps) – DPC 1024-QAM	
BER = 1E-3	74.1
BER = 1E-5	22.9
PAPR	
DPC 16-QAM	5.0
DPC 64-QAM	4.67
DPC 256-QAM	4.79
DPC 1024-QAM	4.91

In this line, given that the PU should operate as in the absence of interference, with no performance degradation, we first propose a scheme based on the trellis shaped DPC encoder for CU. In this sense, some design improvements are discussed in order to overcome the so-called precoding losses. Techniques are detailed for both the encoder and decoder, taking into account the satellite scenario and the trade-off between power efficiency and complexity.

Thereafter, this paper investigates the feasibility of a low data rate secondary service transmission over a primary user infrastructure. A realistic scenario is presented and the previously discussed techniques are implemented to resolve the interference of both links. As a result, we obtained the same performance for the PU as in absence of

the CU operation (AWGN channel). Also, concerning the feasibility analysis, we fulfill the minimum supported bit rate of 16 and 28 kbps (*w.r.t.* BER of 10^{-5} and 10^{-3}, respectively) and the maximum of 22 and 74 kbps (*w.r.t.* BER of 10^{-5} and 10^{-3}, respectively) for the secondary service, where this range is variable according to the implemented scheme. We emphasize that this achievable bit rate is suitable for most of M2M applications.

As a drawback, we point out an increase of the output power of the satellite due to the intrinsic signal correlation presented at the superposition technique. In the described scenario, the total power transmitted is about 32 dBm instead of the previously specified 30 dBm. In this case, the transmitted antenna should be properly designed in order to consider this output power.

Finally, we would recommend further research concerning the following topics: (i) an investigation of the feasibility of a secondary service transmission for GEO multibeam satellite scenario; (ii) the development of a proof of concept by means of SDR implementation; (iii) the analysis of the satellite impairments effects on DPC schemes; (iv) the control of PAPR by shaping operation, typically on satellite communication.

Abbreviations
5G: Fifth generation; AWGN: Additive white Gaussian noise; BER: Bit error rate; CNR: Carrier-to-noise ratio; COTS: Commercial off-the-shelf parts; CR: Cognitive radio; CU: Cognitive user; DPC: Dirty paper coding; DVB-S2X: Digital Video Broadcasting - Satellite - Second-Generation Extension; EESS: Earth exploration satellite service; FEC: Forward error correction code; GEO: Geosynchronous equatorial orbit; HPA: High power amplifier; IoT: Internet of Things; ISI: Inter-symbol interference; ITU-RR: International Telecommunication Union Radio Regulation; LEO: Low Earth orbit; M2M: machine-to-machine; MEO: medium Earth orbit; PAPR: peak-to-average power ratio; PIP: Partial interference presubtraction; PU: Primary user; QAM: Quadrature amplitude modulation; QPSK: Quadrature phase-shift keying; SDR: Software defined radio; SINR: Signal-to-interference-plus-noise ratio; SNR: Signal-to-noise ratio; TCM: Trellis-coded modulation; THP: Tomlinson-Harashima precoding; TP: Trellis precoding; TS: Trellis shaping

Author details
[1] IMT Atlantique, LabSTICC, DEOS, 10 Avenue Edouard Belin, BP 44004, 31028 TOULOUSE CEDEX 4 Toulouse, France. [2] National Institute for Space Research, INPE, Av dos Astronautas, 1758 -. Jardim da Granja, 12227-010 São José dos Campos, Sao Paulo, Brazil. [3] Airbus Defence and Space, Toulouse, France.

References
1. D. Minoli, *Innovations in Satellite Communications and Satellite Technology: The Industry Implications of DVB-S2X, High Throughput Satellites, Ultra HD, M2M, and IP*. (Wiley, 2015)
2. Z. Qu, G. Zhang, H. Cao, J. Xie, LEO satellite constellation for internet of things. IEEE Access. **5**, 18391–18401 (2017)
3. ITU-R, Emerging trends in 5G/IMT2020. Geneva Mission Brief. Ser. (2016). https://www.itu.int/en/membership/documents/missions/gva-mission-briefing-5g-28sept2016.pdf. Accessed 27 Jan 2020
4. Huawei, 55G Network Architecture: A High-Level Perspective. White Paper. Huawei Technologies Co, Shenzhen (2016). https://www.huawei.com/minisite/5g/img/5G_Network_Architecture_A_High-Level_Perspective_en.pdf. Accessed 27 Jan 2020
5. Regulations, ITU Radio, and CHAPTER. II-Frequencies, ARTICLE 5 Frequency allocations. Section IV-Table of Frequency Allocations (2016)
6. DVB, *DVB-Doc A-172" White Paper on the use of DVB-S2X for DTH applications DSNG and Professional Services Broadband Interactive Services and VL-SNR applications*. (ESA, Paris, 2015)
7. United States Department of Commerce, United States radio spectrum frequency allocations chart (2016). https://www.ntia.doc.gov/files/ntia/publications/january_2016_spectrum_wall_chart.pdf. Accessed 27 Jan 2020
8. Y. Han, E. Ekici, H. Kremo, O. Altintas, Spectrum sharing methods for the coexistence of multiple rf systems: A survey. Ad Hoc Netw. **53**, 53–78 (2016)
9. K. Patil, K. Skouby, A. Chandra, R. Prasad, in *2011 The 14th International Symposium on Wireless Personal Multimedia Communications (WPMC)*. Spectrum occupancy statistics in the context of cognitive radio (IEEE, 2011), pp. 1–5
10. M. R. Maheshwarappa, M. Bowyer, C. P. Bridges, in *IEEE Aerospace Conference*. Software defined radio (SDR) architecture to support multi-satellite communications (IEEE, 2015). https://doi.org/10.1109/aero.2015.7119186
11. G. Cocco, T. De Cola, M. Angelone, Z. Katona, S. Erl, Radio resource management optimization of flexible satellite payloads for DVB-S2 systems. IEEE Trans. Broadcast. **64**(2), 266–280 (2017). https://doi.org/10.1109/tbc.2017.2755263
12. N. Porecki, G. Thomas, A. Warburton, N. Wheatley, N. Metzger, in *Proc. 19th Ka Broadband Commun., Navigat. Earth Observat. Conf.* Flexible payload technologies for optimising Ka-band payloads to meet future business needs, (2013), pp. 1–7
13. S. K. Sharma, S. Chatzinotas, B. Ottersten, *Cognitive radio techniques for satellite communication systems* (IEEE, 2013). https://doi.org/10.1109/vtcfall.2013.6692139
14. E. Biglieri, in *2012 IEEE First AESS European Conference on Satellite Telecommunications (ESTEL)*. An overview of cognitive radio for satellite communications (IEEE, 2012). https://doi.org/10.1109/estel.2012.6400078
15. M. Álvarez-Díaz, M. Neri, C. Mosquera, G. Corazza, in *2006 International Workshop on Satellite and Space Communications*. Trellis shaping techniques for satellite telecommunication systems (IEEE, 2006). https://doi.org/10.1109/iwssc.2006.256012
16. A. Jovicic, P. Viswanath, Cognitive radio: an information-theoretic perspective. IEEE Trans. Inf. Theory. **55**(9), 3945–3958 (2009)
17. A. J. Goldsmith, S. A. Jafar, I. Maric, S. Srinivasa, Breaking spectrum gridlock with cognitive radios: An information theoretic perspective. Proc. IEEE. **97**(5), 894–914 (2009)
18. M. Costa, Writing on dirty paper (corresp.) IEEE Trans. Inf. Theory. **29**(3), 439–441 (1983)
19. U. Erez, S. Shamai, R. Zamir, Capacity and lattice strategies for canceling known interference. IEEE Trans. Inf. Theory. **51**(11), 3820–3833 (2005)
20. M. V. Eyuboglu, G. D. Forney, Trellis precoding: Combined coding, precoding and shaping for intersymbol interference channels. IEEE Trans. Inf. Theory. **38**(2), 301–314 (1992)
21. G. Forney, Trellis shaping. IEEE Trans. Inf. Theory. **38**(2), 281–300 (1992)
22. W. Yu, D. P. Varodayan, J. M. Cioffi, Trellis and convolutional precoding for transmitter-based interference presubtraction. IEEE Trans. Commun. **53**(7), 1220–1230
23. Y. Sun, W. Xu, J. Lin, *Trellis shaping based dirty paper coding scheme for the overlay cognitive radio channel* (IEEE, 2014). https://doi.org/10.1109/pimrc.2014.7136456

24. L. B. C. da Silva, T. Benaddi, L. Franck, in *2018 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*. Cognitive radio overlay paradigm towards satellite communications (IEEE, 2018). https://doi.org/10.1109/blackseacom.2018.8433717

25. L. B. C. da Silva, T. Benaddi, L. Franck, in *2018 9th Advanced Satellite Multimedia Systems Conference and the 15th Signal Processing for Space Communications Workshop (ASMS/SPSC)*. A design method of cognitive overlay links for satellite communications (IEEE, 2018). https://doi.org/10.1109/asms-spsc.2018.8510739

26. L. B. C. da Silva, T. Benaddi, L. Franck, in *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*. On the feasibility of a secondary service transmission over an existent satellite infrastructure (Springer, 2019), pp. 154–167. https://doi.org/10.1007/978-3-030-25748-4_12

27. ISIS - Innovative Solutions In Space B.V., ISIS high data rate S-band transmitter specifications. https://www.isispace.nl/wp-content/uploads/2016/02/ISIS-Communicationsystems-Brochure-v2-compressed.pdf. Accessed 27 Jan 2020

28. Secretariat, E. C. S. S. ECSS-M-ST-10C Space Project Management-Project Planning and Implementation (ESA-ESTEC Requirements & Standards Division, Noordwijk, 2009)

29. R. F. Fischer. Precoding and signal shaping for digital transmission (Wiley, 2005)

30. S. Gelfand, Coding for channel with random parameters. Probl. Contr. and Inf. Theory. **9**(1), 19–31 (1980)

31. J. G. Proakis, M. Salehi, Digital communications, vol. 4 (2001)

32. G. Ungerboeck, Channel coding with multilevel/phase signals. IEEE Trans. Inf. Theor. **28**(1), 55–67 (1982)

33. G. Ungerboeck, Trellis-coded modulation with redundant signal sets part I: Introduction. IEEE Commun. Mag. **25**(2), 5–11 (1987)

34. G. Ungerboeck, Trellis-coded modulation with redundant signal sets part II: State of the art. IEEE Commun. Mag. **25**(2), 12–21 (1987)

35. G. D. Forney, L.-F. Wei, Multidimensional constellations. I, introduction, figures of merit, and generalized cross constellations. IEEE J. Sel. Areas Commun. **7**(6), 877–892 (1989)

36. M. Tomlinson, New automatic equaliser employing modulo arithmetic. Electron. Lett. **7**(5), 138–139 (1971)

37. H. Harashima, H. Miyakawa, Matched-transmission technique for channels with intersymbol interference. IEEE Trans. Commun. **20**(4), 774–780 (1972)

38. G. D. Forney, M. V. Eyuboglu, Combined equalization and coding using precoding. IEEE Commun. Mag. **29**(12), 25–34 (1991)

39. G. Maral, M. Bousquet, *Satellite communications systems: systems, techniques and technology*. (Wiley, 2011)

40. D. Barbarić, J. Vuković, D. Babic, in *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. Link budget analysis for a proposed cubesat earth observation mission (IEEE, 2018). https://doi.org/10.23919/mipro.2018.8400026

41. M. Arias, F. Aguado, in *ITU Symposium and Workshop on small satellite regulation and communication systems*. Small satellite link budget calculation, (2016). https://www.itu.int/en/ITU-R/space/workshops/2016-smallsat/Documents/Link_budget_uvigo.pdf. Accessed 27 Jan 2020

Publisher's Note

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