ELSEVIER

CrossMark

Maximizing the link throughput between smart meters and aggregators as secondary users under power and outage constraints

Pedro H.J. Nardelli [a,*], Mauricio de Castro Tomé [a], Hirley Alves [a], Carlos H.M. de Lima [b], Matti Latva-aho [a]

[a] Centre for Wireless Communications (CWC), University of Oulu, P.O. Box 4500, FI-90014, Finland
[b] São Paulo State University, Avenida Doutor Octávio da Silva Bastos, 2439, Bairro Jardim Nova São João 13874–149— São João da Boa Vista-SP, Brazil

ARTICLE INFO

ABSTRACT

This paper assesses the communication link from smart meters to aggregators as (unlicensed) secondary users that transmit their data over the (licensed) primary uplink channel. The proposed scenario assumes: (i) meters' and aggregators' positions are fixed so highly directional antennas are employed, (ii) secondary users transmit with limited power in relation to the primary, (iii) meters' transmissions are coordinated to avoid packet collisions, and (iv) the secondary links' robustness is guaranteed by an outage constraint. Under these assumptions, the interference caused by secondary users in both primary (base-stations) and other secondary users can be neglected. As unlicensed users, however, meter–aggregator links do experience interference from the mobile users of the primary network, whose positions and traffic activity are unknown. To cope with this uncertainty, we model the mobile users spatial distribution as a Poisson point process. We then derive a closed-form solution for the maximum achievable throughput with respect to a reference secondary link subject to transmit power and outage constraints. Our numerical results illustrate the effects of such constraints on the optimal throughput, evincing that more frequent outage events improve the system performance in the scenario under study. We also show that relatively high outage probabilities have little effect on the reconstruction of the average power demand curve that is transmitted from the smart meter to the aggregator.

1. Introduction

Over the last few years, cognitive radios have appeared as the solution for more effective use of the frequency spectrum [1,2]. Following the concept proposed by Haykin [3], the radio nodes should understand their environment to establish a wireless network "(…) with two primary objectives in mind: highly reliable communication whenever and wherever needed; efficient utilization of the radio spectrum." Owing to their cognitive ability, radios would be then capable of sensing the environment to decide on their transmission.

An interesting approach to the cognitive radio concept is the so-called spectrum sharing [4], where unlicensed – secondary – users want to transmit some information without disturbing the licensed – primary – users over the same frequency band. Secondary users then need to sense the spectrum usage to decide about their transmissions. The transmission occurs if the channel is sensed free, which depends on the primary user activity [5]. Otherwise, the secondary

---

* Corresponding author. Tel.: +358 505 210 830.

E-mail addresses: nardelli@ee.oulu.fi (P.H.J. Nardelli), mauricio.decastrotome@oulu.fi (M. de Castro Tomé), halves@ee.oulu.fi (H. Alves), carlos.lima@sjbv.unesp.br (C.H.M. de Lima), matla@ee.oulu.fi (M. Latva-aho).

user either searches for a different band or postpones its transmission.

Although the basic idea is simple, both analysis and implementation are challenging due to interactive dynamics of the decision-making procedures [6]. The solution gets even harder when we consider different possible applications and their specific requirements, yielding no universal solution [7]. Notwithstanding, the idea of spectrum sharing is now widespread in both new generations of cellular system (allowing for co-existence of macro- and small-cells as in [8]) and different sensor applications (e.g. [9–11]).

In this paper, we focus on the specific application of spectrum sharing in the deployment of part of the communication network in the modern electric power grids [12,13] – the so-called smart grids. As one would expect, the different applications of the electric power grid have different requirements from the communication network perspective [14]. For instance, control operations in the high-voltage grid must be close to real-time (scale of milliseconds) so the communication link must have both extremely low latency and very high reliability (greater than 99.9%). In contrast, remote reading of meters would allow for less stringent communication requirements (latency of minutes and a reliability of 98%). As a consequence, the communication network design should be fine-tuned with the specific power grid application and its needs. An informative survey about the different applications and requirements is found in [15].

Looking specifically at how to implement cognitive radio for smart grid applications, the authors in [16] presented an interesting hardware study case using software defined radio in a micro-grid testbed. In [17], a Lyapunov-drift framework was proposed to differentiate traffic priorities and then use the cognitive radio strategy called "dynamic spectrum access" to improve the communications in smart grids. Additionally, Khan et al. [18] provide a recent survey on how to combine cognitive radio approach into smart grid scenarios.

In our case, the authors in [19] and [20] introduce the idea of employing a spectrum sharing scheme within households where home appliances use unlicensed channel bands to build a home area network (HAN). In neighborhood area networks (NANs), different scenarios may be considered for the smart meters: (i) licensed users within the cellular systems directly connected with the distribution operator [21], (ii) users of unlicensed bands of unused TV frequencies (TV white spaces) [22], (iii) secondary users of licensed band [19,20] and (iv) hybrid licensed-unlicensed users [20]. When using solutions (i) and (iv), the smart grid elements are subscribers (licensed users) of cellular networks. In this case, although quality of service shall be guaranteed, they may involve prohibitive high costs and a dramatic increase in data traffic [21], making them not so attractive for electricity providers if compared to (ii) and (iii). And yet, the latter solutions will have a strong dependence on specific country legislation.

Despite such drawbacks, these solutions are interesting and related to ours. We believe, however, that the analysis provided by the aforementioned works are focused on network-layer considerations without a dedicated performance study involving specific topological considerations as node positions and mobility. Herein, we attempt to cope with such limitations by studying a novel application of spectrum sharing technique for distribution grids considering that some elements are fixed and others are mobile.

In specific terms, a spectrum sharing scheme (e.g [4]) will be assumed such that secondary users transmit in the up-link channel of the cellular system. As meters and aggregators are generally static nodes, their communication link may be built using directional antennas (e.g. [23]) while respecting a transmit power constraint. In this way, the harmful effects caused by the secondary user transmissions on the primary users, as well as on other secondary users, due to co-channel interference are limited. By properly designing the antenna beamforming and setting the power constraint, the probability that the secondary users interfere in the primary transmissions would be low as far as directional antennas with limited power have a restrict radiation pattern. Besides, by using the up-link channel, the cellular base-stations are disturbed by the secondary transmissions and, due to their more robust reception procedures [24], the interference can be further mitigated.

On the other hand, the secondary users do experience the interference caused by the primary users. Since the up-link is employed by the former, the interference normally comes from mobile devices, whose positions and traffic are unknown. To account for such uncertainties, the interferers' spatial distribution and traffic characteristics will be modeled using point process theory [25]. Thereby, it is possible to derive closed-form expressions for important performance metrics of wireless systems as outage probability and link throughput. It is worth mentioning that the communication between different smart meters and their respective aggregator is coordinated so they do not interfere to each other.

Our goal in this work is to optimize the link throughput (which is defined as the spectral efficiency, given in bits per second per hertz, times the probability that the packet is successfully decoded by the aggregator) of the secondary link under power (to not affect the primary users) and outage (secondary link reliability) constraints. Using a similar optimization procedure as in [26–28], we find the signal-to-interference ratio threshold and the transmit power employed by the secondary link so as to maximize its throughput while respecting the imposed constraints.

Then, we found a closed-form equation for the optimal link throughput as a non-linear function of the system constraints and the density of interferers, as well as other system parameters. Interestingly, our results show that frequent outage situations have an unexpected positive effect on the system performance for the scenario of interest (maximum value of outage constraint is 25%). We use actual data from the average power demand of a household (obtained from "The Reference Energy Disaggregation Data Set" database [29,30]) to show that relatively high outage probabilities do not lead to a poor signal reconstruction if the information is sampled and sent periodically, for example, every 15 min. All in all, our main contribution and novelty are the theoretical evidences that up-link spectrum sharing is a good candidate for deploying the communication network in the distribution electricity grids.

The rest of this paper is divided as follows. Section 2 explains the scenario under analysis, justifying our assumptions and presenting the optimization problem to be solved. In Section 3, we focus on the solution of the optimization

problem and illustrate how the system performance changes with the configuration parameters. Section 4 shows how the outage events affect the reconstruction of the average power demand curve. In Section 5, we discuss some implications of our theoretical results, indicating how they might be used in actual deployments. Section 6 concludes this paper.

2. System model

In this section, we describe the basic assumptions used to build our model and their implications. The assumptions are stated as follows:

- **Assumption 1**: Spectrum-sharing scenario where licensed (primary) and unlicensed (secondary) users share the frequency bands allocated to the up-link channel.
- **Assumption 2**: Primary link is established between static cellular base-stations and mobile users. Secondary terminals are smart meters that need to forward data to a given aggregator through the uplink channel and their positions are fixed.
- **Assumption 3**: The smart meters transmit with limited power W_s such that $W_s \leq W_{max}$ where W_{max} is the maximum power allowed for the secondary users (which can be seen as an imposition from the primary network).
- **Assumption 4**: Smart meters associated with the same aggregator are able to perfectly coordinate their transmissions using time scheduling.

In this case, Assumption 2 indicates the possibility of employing directional antennas in the secondary links as far as their positions are fixed. Orientation errors as defined in [23] can be then completely avoided when deploying the secondary network, making highly directional antennas worth. In its turn, Assumption 3 imposes the maximum range that the signal transmitted by the smart meters can reach. Putting all together, the radiation pattern generated by the secondary transmission can be seen as a line segment starting in the smart meter, passing through the aggregator and ending in a point related to W_{max}.

Let us now look at the interference related to the proposed spectrum sharing. Assumptions 1, 2 and 4 indicate the co-channel interference occurs: (i) from smart meters to cellular base-stations, (ii) from smart meters to aggregators that they are not associated, and (iii) from mobile users to aggregators. From the implications discussed in the previous paragraph, the cases (i) and (ii) can be neglected by designing the specific locations when the deploying either the secondary or primary networks. Even if the positions are considered random in two-dimensions, the chance of having a base-station or an aggregator in the line segment related to the smart meter transmitted signal approaches zero, which further indicated that such cases should not be considered.

Hence, only case (iii) is relevant for our analysis. To evaluate its impact in the system performance, we first need to model uncertainty of the mobile users' positions and traffic activity. We assume here a Poisson field of interferers [31] such that the interferer nodes are distributed over an infinite plane following a 2-dimensional Poisson point process Φ with density λ, given in interferers per square-meter. The wireless channel model employed in this paper consists of two components: one related to the distance-dependent path-loss such that the received power decays with the distance and other related to fast-fading [25]. The received power at the node of interest can be computed as $g_i r_i^{-\alpha}$, where r_i is the distance between the reference receiver and the ith node, g_i is the channel gain between them, and $\alpha > 2$ the path-loss exponent.

Hereafter, we focus our analysis on a reference smart meter-aggregator link, as shown in Fig. 1. During transmissions intervals, we assume that the interferers' positions and the channel gains do not change. We also consider an interference-limited scenario wherein the noise effects can be neglected. As pointed in [32], the inclusion of the noise power leads to a more complex analysis without providing any significant qualitative difference.

If the primary users are equipped with omni-directional antennas and transmit with the same fixed power W_p, the
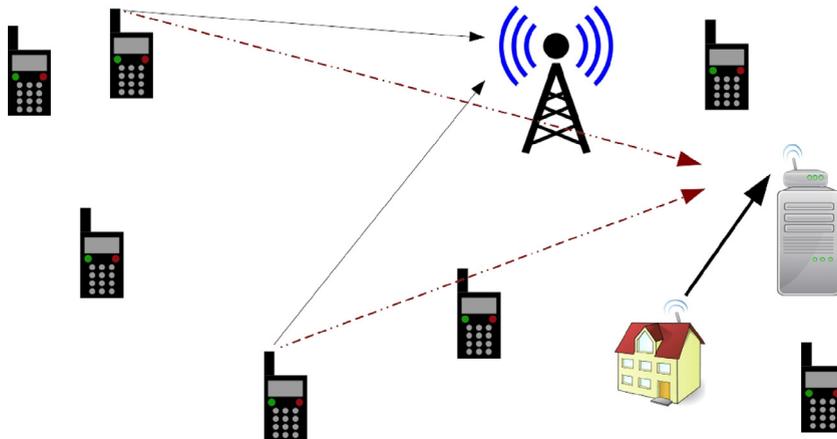
Fig. 1. An illustration of the proposed scenario, where primary and secondary users share the up-link channel. The reference smart meter (secondary transmitter) is depicted by the house, the aggregator (secondary receiver) by the the CPU , the handsets are the mobile primary users (interferers to the aggregator) and the big antenna is the cellular base-station. As the smart meter uses directional antenna with limited transmit power (bold arrow), its interference towards the base-station can be ignored. The thin black arrows represent the primary users' desired signal, while the red ones their interference towards the aggregator. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

signal-to-interference ratio (SIR) at the aggregator can be computed as [25]:

$$\text{SIR}_0 = \frac{W_s g_0 r_0^{-\alpha}}{W_p \sum_{i \in \Phi} g_i r_i^{-\alpha}}, \tag{1}$$

where the index 0 denotes the reference link.

We assume that the reference link employs both point-to-point Gaussian codes and the interference-as-noise decoding rule [28,33] so that a spectral efficiency of $\log_2(1 + \beta)$ in bits/s/Hz is achieved only if the SIR is greater than β. In this case, the probability P_{suc} that a packet is successfully decoded by the aggregator is the probability that $\text{SIR}_0 > \beta$. Then, an outage event happens with probability $1 - P_{\text{suc}}$. In our scenario, retransmissions are not allowed so that the information contained in packets received in outage is lost.

To compute P_{suc}, we assume quasi-static channel gains (squared envelops) g that are independent and identically distributed exponential random variables (Rayleigh fading), and also a dynamic topology where interferers' positions change every transmission interval. Therefore, every transmission attempt can be viewed as a different realization of the point processes Φ and the channel gains g. We consider here that the distance between the reference meter and aggregator is known and has a fixed value $r_0 = d$. Finally, the success probability is computed as [25]:

$$P_{\text{suc}} = e^{-\lambda \kappa \pi d^2 \beta^{2/\alpha} \left(\frac{W_p}{W_s} \right)^{2/\alpha}}, \tag{2}$$

where $\kappa = \Gamma(1 + 2/\alpha)\Gamma(1 - 2/\alpha)$ with $\Gamma(\,\cdot\,)$ being the Gamma function. It is also worth noting that the primary users might employ different power control strategies, but it will not affect the qualitative results of our analysis due to the stochastic geometry approach employed (e.g. [34]).

We are now ready to define the performance metric and carry out the optimization problem under consideration.

Definition 1 (Link throughput). The throughput T of the reference link using the system model described in this section is defined as:

$$T = \log(1 + \beta) \, P_{\text{suc}} = \log_2(1 + \beta) \, e^{-\lambda \kappa \pi d^2 \beta^{2/\alpha} \left(\frac{W_p}{W_s} \right)^{2/\alpha}}. \tag{3}$$

Our goal in this paper is to find the setting of parameters for the secondary users to maximize their link throughput T while respecting the imposed power limit and outage constraints. In our case, the variables in hand are the coding rate β and the transmit power W_s of the reference link. Mathematically, we have the following problem:

$$\max_{(\beta, W_s)} \quad \log_2(1 + \beta) \, e^{-\lambda \kappa \pi d^2 \beta^{2/\alpha} \left(\frac{W_p}{W_s} \right)^{2/\alpha}}$$
$$\text{s.t.} \quad W_s \leq W_{\max} \tag{4}$$
$$\qquad 1 - P_{\text{suc}} \leq \epsilon$$

where ϵ is the maximum acceptable outage probability, reflecting the reference link robustness.

3. Maximum throughput under power and outage constraints

In this section, we solve the optimization problem previously stated. We then provide some numerical results to illustrate how the constraints imposed to our smart meter–aggregator reference link will affect the maximum achievable throughput. Before we start, we still need to present a Lemma that tells us how the throughput behaves as a function of the secondary transmit power W_s and the SIR threshold β when no constraint is considered.

Lemma 1. *Let us consider the throughput equation, given by (3), as a function of the variables $W_s > 0$ and $\beta > 0$, i.e. $T = f(W_s, \beta)$. The function f is monotonically crescent in respect to W_s, and it is concave in respect to β if $\partial^2 T / \partial \beta^2 < 0$.*

Outline of proof. The proof of this Lemma is straightforward from the analysis of the function in terms of the (strictly positive) variables W_s and β, and knowing that the function T is twice differentiable in terms of β. Any increase in W_s leads to an increase in the exponential term of (3) and then in T. Increasing β, on the other hand, has a two-fold effect: it increases the logarithmic term while decreases the exponential one. The function T, however, is not always concave in relation to $\beta > 0$; nevertheless, in the region that $\partial^2 T / \partial \beta^2 < 0$, T is concave. □

Since T is concave for some values of β, we may try to find the value of β that leads to the maximum T. This is shown in the next Lemma.

Lemma 2. *Let β_{un}^* denote the value of β that maximizes T assuming that β_{un}^* is in the region where $\partial^2 T / \partial \beta^2 < 0$. Then, β_{un}^* is the solution of the following (transcendental) equation:*

$$\alpha \beta = k \, \beta^{2/\alpha} \, (1 + \beta) \, \ln(1 + \beta), \tag{5}$$

where $k = 2\lambda \kappa \pi d^2 (W_p / W_s)^{2/\alpha}$.

Outline of proof. Let us first consider that β_{un}^* is in the region where the inequality $\partial^2 T / \partial \beta^2 < 0$ holds. Then, from Lemma 1, the value of β that maximizes T is the solution of the derivative equation $\partial T / \partial \beta = 0$. By doing so in (3), we end up in (5), which has no closed-form solution. If a solution does not exist, then our initial assumption the inequality $\partial^2 T / \partial \beta^2 < 0$ does not hold and β_{un}^* cannot be obtained. □

Nonetheless, in the cases of interest, (5) has a solution and therefore β_{un}^* is in the region where $\partial^2 T / \partial \beta^2 < 0$ as we shall see next. It is noteworthy that even though (5) does not have a closed-form solution, it can be easily evaluated numerically through standard mathematical software as, for instance, in [35].

Fig. 2 exemplifies the behavior of the link throughput T with the transmit power W_s and the threshold β. The curve behaves as predicted by Lemmas 1 and 2. Intuitively, when all parameters are kept the same, it is advantageous for the smart meter to increase its transmit power so that the SIR experienced by the aggregator tends to increase. This behavior, although individually optimal, is not good for the other links as transmitting with more power would increase the interference level throughout the network (refer to [27] for a deeper assessment of an optimization problem wherein the optimal individual solution can be socially harmful). For this reason, the power constraint that our problem assumes is needed, as we will discuss later.

When analyzing the effects of the threshold β, there is a trade-off involved. An increase of β leads to a more efficient transmission where more bits/s/Hz can be transmitted
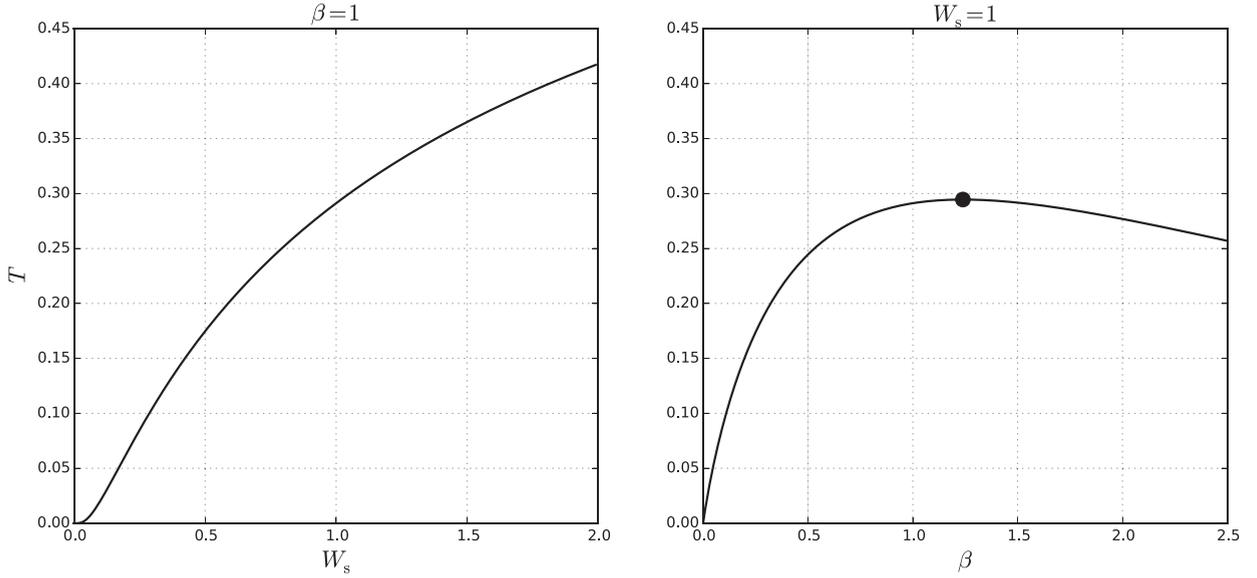
Fig. 2. Link throughput T given in (3) as a function of the transmit power W_s (left) and the SIR threshold β (right) for $\alpha = 4$, $d = 1$, $\lambda = 0.25$ and $W_p = 1$. The black dot in the right plot is the optimal operating point predicted by Lemma 2.

in the same message. This gain, however, comes at expense of more frequent outage events, which in turn decreases the link throughput. From Lemma 2, the value of β that leads to the optimal operating point can be found by numerically solving (5). In our case, we use the function fsolve( ) from the Python library SciPy [35].

We are now almost ready to present the main result of this paper, which is the solution of the constrained optimization problem presented in the previous section. But first, we still need to state a lemma about the relation between the optimal values of the constrained and unconstrained SIR thresholds β.

Lemma 3. *Let β^* denote the value of β that leads to the maximum constrained throughput given in (4) and β_{un}^*, given in Lemma 2, denote the value that optimizes the unconstrained throughput. Then:*

$$1 - e^{-\frac{\alpha \beta_{un}^*}{2(1+\beta_{un}^*)\ln(1+\beta_{un}^*)}} > \epsilon \implies \beta^* < \beta_{un}^*. \tag{6}$$

Proof. In the optimal unconstrained operating point β_{un}^* given in Lemma 2, the following equality must hold:

$$\beta^{2/\alpha} = \frac{\alpha \beta}{k(1+\beta)\ln(1+\beta)}. \tag{7}$$

Then, although we cannot analytically compute the actual value of β_{un}^*, we do know that the outage probability $1 - P_{suc}$, given in (2), related to it will be given by: $1 - e^{-\frac{\alpha \beta_{un}^*}{2(1+\beta_{un}^*)\ln(1+\beta_{un}^*)}}$, which is a monotonically decreasing function of β_{un}^*.

In this case, if that probability is smaller than the constraint ϵ, then the constrained optimal threshold is the unconstrained one, i.e. $\beta^* = \beta_{un}^*$. Otherwise, the optimal constrained threshold β^* must be smaller than β_{un}^* since the outage probability is a decreasing function of β_{un}^*. □

Proposition 1. *Let us assume that the pair (W_s^*, β^*) is the solution of the optimization problem given by (4). If $1 -$*

$e^{-\frac{\alpha \beta_{un}^*}{2(1+\beta_{un}^*)\log(1+\beta_{un}^*)}} > \epsilon$, *the pair (W_s^*, β^*) is computed as:*

$$W_s^* = W_{max}, \tag{8}$$

$$\beta^* = \frac{W_{max}}{W_p}\left(-\frac{\ln(1-\epsilon)}{\lambda \kappa \pi d^2}\right)^{\alpha/2}. \tag{9}$$

The optimal throughput T^* is then:

$$T^* = (1 - \epsilon) \times \log_2\left(1 + \frac{W_{max}}{W_p}\left(-\frac{\ln(1-\epsilon)}{\lambda \kappa \pi d^2}\right)^{\alpha/2}\right). \tag{10}$$

Proof. Let us start by considering the variable W_s. From Lemma 1, we know that the throughput is a monotonically increasing function of W_s, regardless of β. Then, W_s must assume its highest possible value: $W_s^* = W_{max}$.

To find β^*, we first use the assumption that $1 - e^{-\frac{\alpha \beta_{un}^*}{2(1+\beta_{un}^*)\ln(1+\beta_{un}^*)}} > \epsilon$. From Lemma 3, the inequality $\beta^* < \beta_{un}^*$ holds. Then, we use the fact that T is also a monotonically crescent function of β in the range $0 < \beta < \beta_{un}^*$ so that β^* should be the highest value that satisfies the inequality $1 - P_{suc} \leq \epsilon$. Manipulating the constraint by assuming $W_s^* = W_{max}$, we obtain $\beta \leq \frac{W_{max}}{W_p}\left(-\frac{\ln(1-\epsilon)}{\lambda \kappa \pi d^2}\right)^{\alpha/2}$. In this case, equality gives β^*. □

Remark 1. This result is only valid if the initial assumption $1 - e^{-\frac{\alpha \beta_{un}^*}{2(1+\beta_{un}^*)\ln(1+\beta_{un}^*)}} > \epsilon$ holds, which is true for the cases of interest, namely $\alpha \in (2, 6]$ and $\epsilon \in (0, 0.25)$. Notice that $\alpha \in (2, 6]$ comprises indoor and outdoor scenarios as well as rural and urban areas [36]. For example, when $\alpha = 4$ and $\beta = 1.24$ (the value of β_{un}^* in the example used in Fig. 2), the outage probability is 0.75. If that inequality does not hold, the optimal solution is $\beta^* = \beta_{un}^*$ and the optimal power W_s^* should be computed accordingly. In other words, the value of
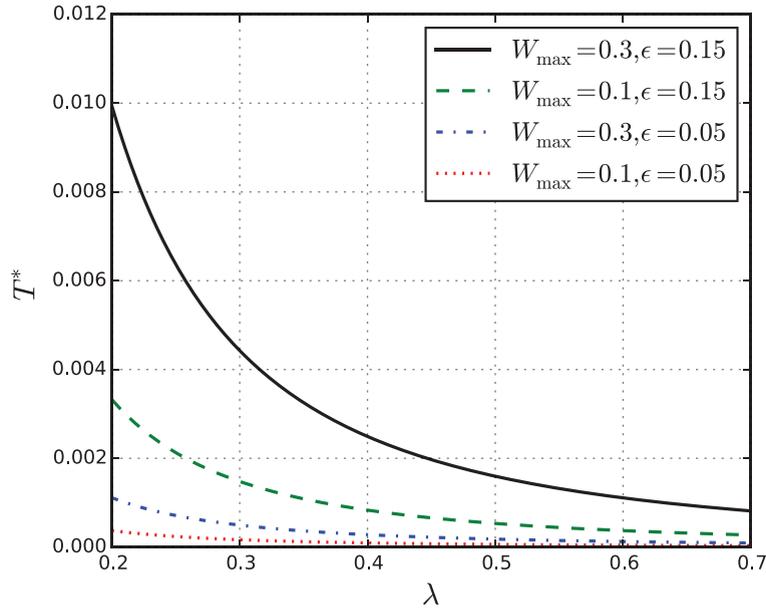
Fig. 3. Maximum link throughput T^* given in (10) as a function of the density of active mobile primary users (interferers) λ for $\alpha = 4$, $d = 1$, $W_p = 1$ and different values of the constraints W_{max} and ϵ.

β^* is kept fixed, while the optimal power W_s^* is the variable used to optimize the link throughput.

Corollary 1. *The optimal throughput T^* can be approximated by:*

$$T^* \approx \frac{(1-\epsilon)\,W_{max}}{\ln(2)\,W_p}\left(-\frac{\ln(1-\epsilon)}{\lambda \kappa \pi d^2}\right)^{\alpha/2}, \tag{11}$$

when $\frac{W_{max}}{W_p}\left(-\frac{\ln(1-\epsilon)}{\lambda \kappa \pi d^2}\right)^{\alpha/2}$ is small.

Remark 2. Corollary 1 holds for the cases under study. For example, when the system variables are set as follows: $\lambda = 0.25$, $\alpha = 4$, $d = 1$, $W_p = 1$, $W_p = 0.5$ and $\epsilon = 0.05$, the term is $\frac{W_{max}}{W_p}\left(-\frac{\ln(1-\epsilon)}{\lambda \kappa \pi d^2}\right)^{\alpha/2} = 0.00086$, therefore the approximation $\ln(1 + x) \approx x$ works well.

Next we will illustrate the analytic results just presented to get more insights on how the maximum link throughput T^* is affected by the activity of the primary users, as well as the system constraints. Fig. 3 shows how the optimal constrained throughput T^* in the smart meter-to-aggregator link behaves in relation to the density λ of active mobile users that interfere in its communication. As expected, increasing the density of interfering nodes decreases the maximum throughput achieved by the reference link, regardless of values of W_{max} and ϵ assumed. It is interesting to see that T^* exponentially decays with λ, which indicates that the secondary link performance is dramatically affected by the primary users' increase of activity.

We can also see from Fig. 3 that the values of the constraints W_{max} and ϵ affect the maximum throughput curves. Higher values of either W_{max} or ϵ leads to higher T^*, when λ is fixed. While the result is intuitive for W_{max} (by increasing the transmission power, we obtain higher SIR and link

throughput), it is not so when the outage constraint ϵ is considered: a less strict constraint leads to higher throughputs.

To better understand those behaviors, we present in Figs. 4 and 5 the maximum throughput T^* versus W_{max} and ϵ, respectively. Fig. 4 shows that the maximum throughput T^* linearly grows with W_{max}, which was predicted by Corollary 1. This means that any relaxation in the power constraint W_{max} provides a linear gain in the secondary link throughput, whose slope is directly defined by the system variables; therefore, a combination between a relatively high outage constraint and low density of interferers provides the best performance.

The behavior of T^* as a function of the outage constraint is more complicated since there is a trade-off involved, as shown in (10) and (11). However, from our assumption that ϵ is a relatively small probability, i.e. $\epsilon \in (0, 0.25)$, then T^* is a (non-linear) crescent function of ϵ. Fig. 5 illustrates this growth, evincing that an increase $\Delta\epsilon$ for smaller values of ϵ leads to lower variation ΔT^* in the maximum throughput. In this case, allowing for more outage events is more advantageous for the link: the spectral efficiency gains obtained by setting higher SIR thresholds dominates the system performance. It is worth reinforcing that this behavior – which is somehow counter-intuitive – is only valid for lower values of $\epsilon \in (0, 0.25)$. If outages are not a constraint in our system, the results of Proposition 1 should be reviewed under the perspective of Lemmas 1 and 2, and Remark 1.

4. Outage events and signal reconstruction

We just showed in the previous section that allowing for more frequent outage events increases the link throughput. However, outage events will affect the information that needs to be sent by the smart meter to the aggregator.
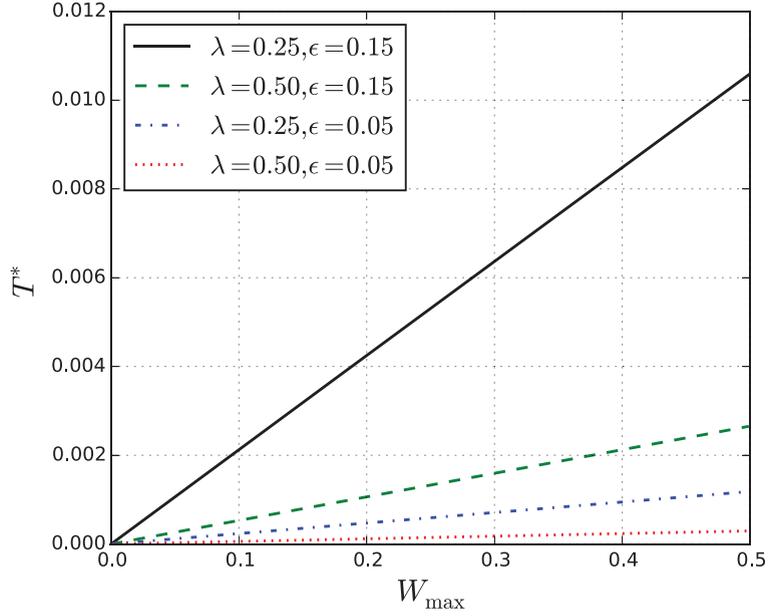
Fig. 4. Maximum link throughput T^* given in (10) as a function of the secondary power constraint W_{max} for $\alpha = 4$, $d = 1$, $W_p = 1$ and different values of λ and ϵ.
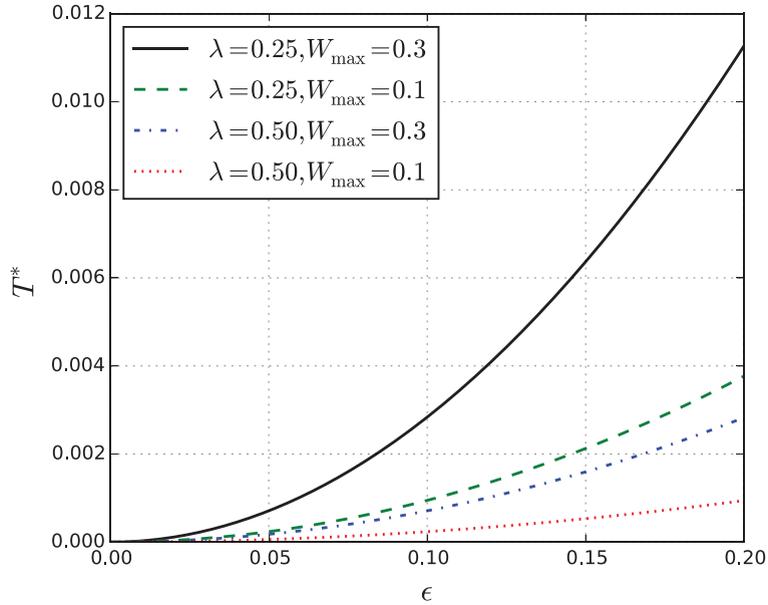
Fig. 5. Maximum link throughput T^* given in (10) as a function of the outage constraint ϵ for $\alpha = 4$, $d = 1$, $W_p = 1$ and different values of λ and W_{max}.

Let $x[n]$ be the discrete signal transmitted by the smart meter and $\hat{x}[n]$ be the signal received by the aggregator, where $n = 1, \ldots, N$ with N being the last sample. We consider the $x[n]$ is the average power demand over a fixed period of time τ. The aggregator needs then to use $\hat{x}[n]$ to reconstruct the signal $x[n]$. We assume here that the signal is reconstructed via linear interpolation between two adjacent points. If the communication is perfect (i.e. $\hat{x}[n] = x[n]$) the interpolation is always between $\hat{x}[k]$ and $\hat{x}[k-1]$, with $k = 2, \ldots, N$.

This, however, is not the case in our model since outage events may occur due to the primary users activity. If

a sample is lost, the aggregator will interpolate the missing value(s) using the latest two received samples. Consider the transmitted sequence: $x[k-2], x[k-1], x[k]$ with $k = 2, \ldots, N$. If the samples $x[k-2]$ and $x[k]$ are successfully received but $x[k-1]$ is not, the reconstruction is based on the linear interpolation of $\hat{x}[k] = x[k]$ and $\hat{x}[k-2] = x[k-2]$. The estimation of the missing point is then $\hat{x}[k-1] = (x[k] + x[k-2])/2$.

To carry out this procedure, we use the "The Reference Energy Disaggregation Data Set" database [29,30] from where we build our signal $x[n]$, which is the 15-minute average power demand over a timespan of 24 h (one day). The
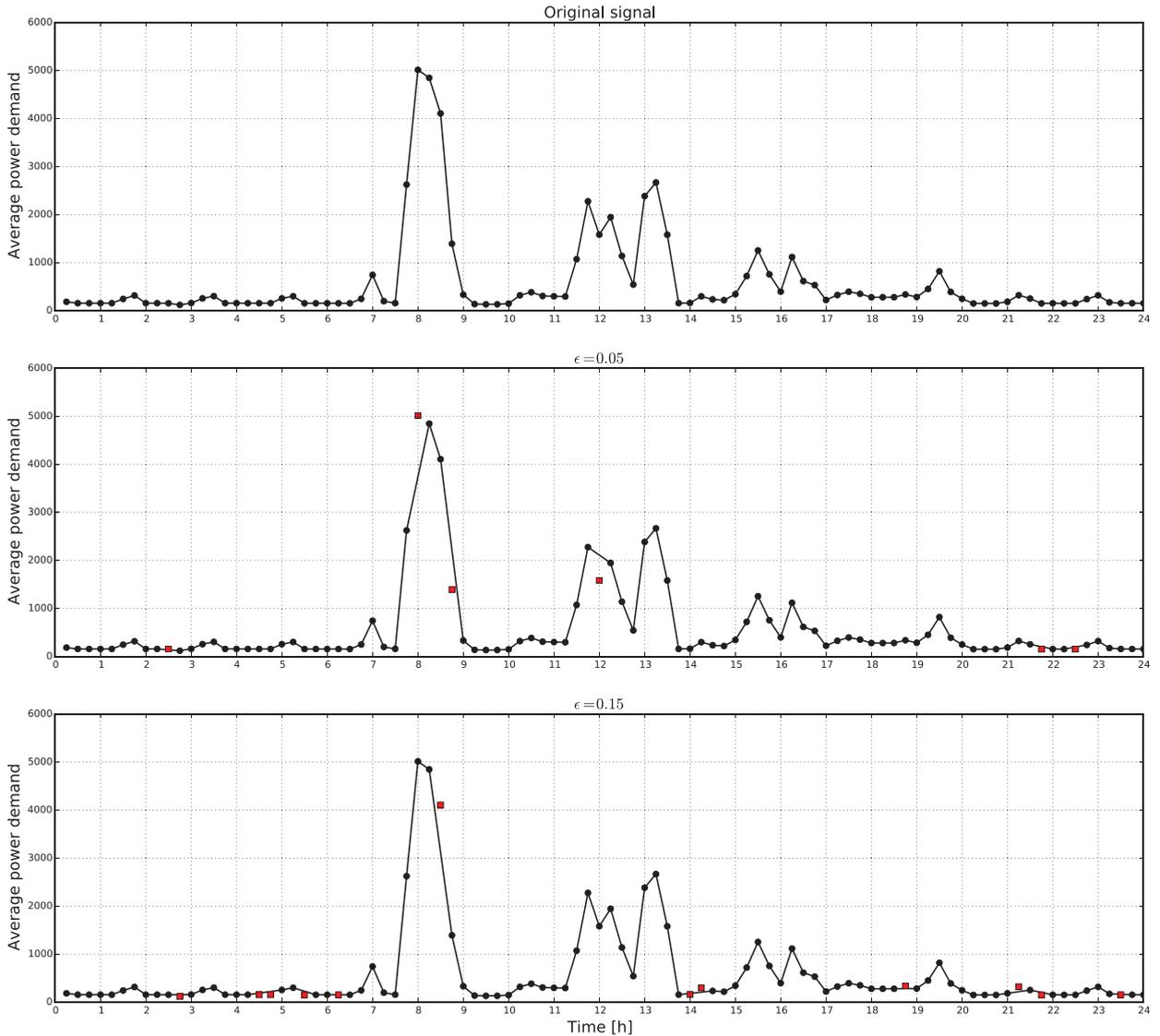
Fig. 6. Average power demand of a house measured in watts during a period of 24 h. The signal is reconstructed in the aggregator as a linear interpolation between two subsequent points. If an outage happens, the point related to the power demand at that time is lost. On the top: perfect reconstruction. At the center: outage probability $\epsilon = 0.05$. On the bottom: outage probability $\epsilon = 0.15$. The red squares in the last two plots are the missing points (samples). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

information from the smart meter is transmitted to the aggregator every 15 min (yielding $\tau = 0.25$ h and $N = 96$ samples), which reconstructs the signal as previously described.

Fig. 6 illustrates the effects of the outage probability on the signal reconstruction by showing the original signal and a snapshot of outputs assuming the constraints $\epsilon = 0.05$ and $\epsilon = 0.15$, where the missing points in the $\epsilon = 0.05$ and $\epsilon = 0.15$ curves are identified by the red squares. At first sight, one cannot notice a big difference between the three signals, and the most noticeable differences are in the $\epsilon = 0.05$ curve (specifically in the "peaks"), not in the $\epsilon = 0.15$ one. This interesting fact happens due to the nature of the signal itself and the randomness of outage events. The power demand signal seems to have a burst nature with a floor level and few peaks, which is related to personal habits when the house is

occupied and people are engaged in activities like cooking or showering.

For example, we can see in Fig. 6 that the peaks are around 8:00 in the morning and noon. These peaks are probably related to people getting ready to work and having lunch. Other than this, specially between 10:00 in the evening and 6:00 in the morning of the following day, we can see that the energy consumption is quite low and constant, most probably related to appliances in stand-by and refrigerator cycles [37]. In this way, most of the samples $x[k]$ will have similar values. If independent and identically distributed erasures occur, the probability that the estimated point approximates the missing one is high.

Consider again the transmitted sequence: $x[k-2]$, $x[k-1]$, $x[k]$ with $k = 2, \ldots, N$ and that the samples $x[k-2]$ and
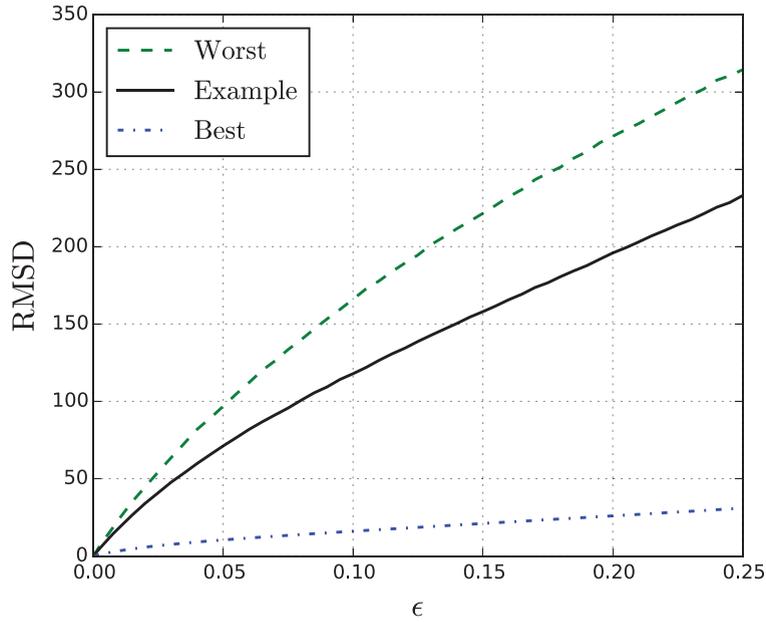
Fig. 7. Root-mean-square deviation (RMSD) versus the outage constraint ϵ. The results are presented for the example given in Fig. 6 together with the best and worst cases in terms of RMSD among 53-household demand data from [30]. The curves are the result of Monte Carlo simulations considering 10^5 realizations for each point.

$x[k]$ are successfully received but $x[k-1]$ is not. Then, the reconstruction is $\hat{x}[k-1] = (x[k] + x[k-2])/2$. As in most of the cases $x[k] \approx x[k-1] \approx x[k-2]$, then $\hat{x}[k-1] = (x[k] + x[k-2])/2 \approx x[k-1]$. However, during the peaks, this does not hold and errors become evident. The snapshot presented for $\epsilon = 0.05$ is an unlucky one as far as the transmissions failure happened in the peak periods. Conversely, although more samples were lost when $\epsilon = 0.15$, they were mostly in the floor-level periods and the signal reconstruction was not affected in this specific snapshot.

To statistically analyze the effects of the outage events in the signal reconstruction, we use the root-mean-square deviation (RMSD) such that the reconstruction error of the $\hat{x}[n]$ is computed as:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (\hat{x}[k] - x[k])^2} . \qquad (12)$$

Fig. 7 presents how the RMSD changes with the maximum allowed outage probability ϵ. As expected, more frequent outage events (indicated by higher ϵ) leads to a worse signal reconstruction (indicated by higher RMSDs). Our results show, however, that ϵ has relatively little effect on reconstructing the signal. Let us first analyze the example presented in Fig. 6 (black curve in Fig. 7). For the highest value considered $\epsilon = 0.25$, the RMSD is 227 watts for a signal with mean of 587 watts and with range $\max x[n] - \min x[n] = 4895$ watts. In this case, the normalized RMSD with respect to the mean is approximately 0.39 while with respect to the range is less than 0.05. Looking back to the signal itself in top plot of Fig. 6, one can see that there are few points that, if erased, would cause a significant distortion on its reconstruction.

In order to have a more robust analysis, we also present in Fig. 7 the best and the worst cases in terms of RMSD among 53 households, whose data is available in [30]. For this scenario, our example is closer to the worst case than to the best one. We further show in Fig. 8 the frequency diagram (histogram) of RMSD for the scenario where $\epsilon = 0.25$. As one can see, most households perform better than the example presented Fig. 6 (RMSD = 227 W); they are in fact much closer to the best case presented in Fig. 7. All in all, these results reinforces even more our argument that, for the scenario under consideration, the signal reconstruction is weakly affected by relatively frequent outage events.

5. Discussions

Throughout the last two sections, we studied a spectrum sharing scenario where a given smart meter–aggregator pair communicates using a licensed cellular uplink channel. For this scenario, we analytically assessed how the system constraints for being a secondary user affect the link throughput and the signal reconstruction. In this section, we start from the presented results to discuss how the proposed spectrum-sharing scheme could be implemented in actual systems.

5.1. Relation between licensed and unlicensed users

In our theoretical model, we assume that the secondary, unlicensed, users do not affect the primary users (base-stations). As discussed before, this can indeed be the case when deploying our strategy in a real system. Since the positions of the smart meters, aggregators and base-stations are fixed, the use of highly directional antennas in the smart meter–aggregator link with low transmit power will decrease the interference level at the base-station, which in
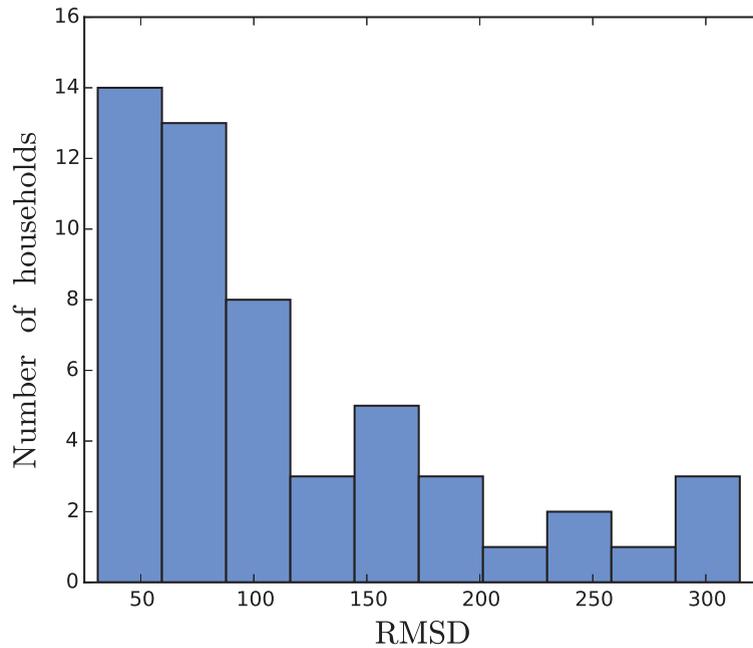
Fig. 8. Frequency diagram of the root-mean-square deviation (RMSD) considering data from 53 households from [30] for $\epsilon = 0.25$ versus the outage constraint ϵ. The results are based on Monte Carlo simulations using 10^5 realizations for each point.

turn still has capabilities of dealing with such residual interference.

Looking at the interference caused by the mobile users in the aggregator, our analysis assume that the density of these nodes are fixed. This, however, will probably not hold because the primary user activity probably would intensify during periods that match with power demand peaks. Then, in a practical scenario, this traffic variations is somehow predicted and the smart meter may adaptively change its operation point either by setting a pre-determined, predicted, density of interferers to optimize its communication link (simpler solution) or by sensing and estimating such a density from time to time (more complex solution).

Another important point that our model can be useful relates to Denial-of-Service (DoS) attacks in the physical layer [38]. In this scenario, a node or a group of nodes intentionally interfering in the aggregator reception could characterize a DoS. From the analysis presented here, we could infer how frequent outage events need to be to affect the signal reconstruction. A dedicate study would be also possible by modeling the attackers as another Poisson point process.

5.2. Outage constraint and link robustness

While the power constraint is required by the secondary link to not interfere in the primary users, the outage constraint is set to guarantee a minimum robustness at the secondary link. Conversely to what one would expect, our results show that allowing for more frequent outage events improves the link throughput due to the contradictory effects of the SIR requirement on the system performance (lower outage constraints lead to higher SIR constraints, which results in higher spectral efficiency, while it decreases the success probability).

The question that arises from this result is *how robust against outage events the secondary link should be?* One can only reply this answer knowing the information that is sent to the aggregator. Our example shows that, if the information to be sent is the average power demand, the signal reconstruction is possible even with relatively loose outage constraints. While this happens due to the nature of the input signal as discussed in the previous section, higher outage probabilities might be not desirable for other kind of signals or if the aggregator should provide some kind of feedback to the smart meter change the power demand behavior, as in strategies of demand-side management [39].

5.3. Power demand signal processing and transmission

The signal presented in Fig. 6 exemplifies a 15-minute sampling interval of the power demand of a household. The signal characteristic indicates that time-based sampling might not be the most efficient way of collecting and then send the data to the aggregator. For instance, by using event-based sampling [40,41], the communication link should be much more robust (i.e. lower values of the outage constraint) since there will be much less redundant data and therefore the loss of any sample will have a more dramatic effect on the signal reconstruction.

Although this is not the focus of the present paper, we would like to mention that there is a trade-off between the sampling strategies and communication. A more efficient way of sampling leads to less points for reconstructing the original signal and vice-versa. If this is the case, the transmission strategies and then the outage constraint should be evaluated in combination with the sampling procedure. In the scenario used here, the time-based sampling generate

redundant information about the signal so that outage events do not have drastic effects on the signal reconstruction.

For this new study, however, it is important to have a stochastic characterization of power demand signal as in [42,43]. By doing so, the signal processing framework would be generalized and adapted to different consumption patterns. This would allow us to study ways to optimize different sampling strategies (time-based, event-based and hybrid) and transmission system designs (coding rate, medium access protocol and retransmissions), while considering the different requirements as reconstruction error and information privacy. Looking at the reconstruction procedure, more advanced strategies like machine learning would probably offer better options to reconstruct the data than linear interpolation, but at expense of higher computational costs [44]. From this perspective, it would be worth evaluating the cost of the sampling-reconstruction and the benefits of link throughput.

In this paper, however, we have chosen to employ a two-dimensional analysis so as to understand the relations between optimal throughput and the reconstruction error itself; our goal is provide the knowledge of what combination of requirements is possible to achieve. For example, an application that requires a very high reliability would imply in a lower throughput. Consequently, a cost-benefit analysis involving those aspects would require more information about the application requirement.

Another different scenario might consider (hybrid) Automatic Repeat Request (ARQ) strategies in order to enhance the communication link, which reduces outage events [26]. This scenario may also include MAC protocols in order to control re-transmissions as well as coordination among nodes. One recent example of a MAC protocol designed in a similar scenario was developed in [45], where the authors introduced a centralized strategy that resorts to a specialized frame structure that support co-existence between a cognitive and primary networks. Along the same line, the idea to include a routing protocol that protects the primary users while meeting the utility requirements of smart grid network was proposed in [46]. Another scenario of interest deals with security and privacy of the transmitted data, since there is a trade-off between security and reliability as pointed out in [47], where the authors resort to information theoretical tools in order to guarantee security already at the physical layer. All in all, combining these more advanced communication techniques to decrease the outage events with a more efficient sampling and enhanced security at the physical layer would be an interesting next step for the present work.

6. Conclusions

This paper assesses a spectrum sharing scenario where smart meters send periodic information to an aggregator over licensed cellular uplink channels. We assume that the secondary link uses directional antennas with limited transmit power so its interference in the primary users can be neglected. Mobile primary users, on the other hand, interfere with the aggregator reception. Modeling the interferers' spatial distribution as a Poisson point process, we analyzed the secondary link throughput, finding then its optimal value under power and outage constraints.

Our results show that relatively high outage constraints surprisingly improve the link throughput for the cases of interest, even though more samples will be lost. It is worth mentioning the fact that smart meter reading application with scheduled intervals requires nowadays a reliability of at least 98% [15, Table 3]; this is a fairly high value if compared to our results. In fact, the discrepancy between the actual requirement and our results was a surprise. We also show that, due to the burst nature of the power demand signal that is transmitted in the smart meter–aggregator link, outage events do not have a dramatic effect in the signal reconstruction in comparison to the perfect transmission.

We plan to study in future works how outage events will affect the signal reconstruction under different sampling strategies. In this way, we plan to build a joint sampling-transmission technique that can improve the system efficiency, as discussed in Section 5.3. We also expect to implement the ideas proposed here in an actual demonstration to verify the validity of our assumptions and proposed optimization strategy.

Acknowledgment

References

[1] I.F. Akyildizet, et al., Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey, Comput. Netw. 50 (13) (2006) 2127–2159.

[2] C. Cormio, K.R. Chowdhury, A survey on MAC protocols for cognitive radio networks, Ad Hoc Netw. 7 (7) (2009) 1315–1329.

[3] S. Haykin, Cognitive radio: brain-empowered wireless communications, IEEE J. Sel. Areas Commun. 23 (2) (2005) 201–220.

[4] J. Peha, Sharing spectrum through spectrum policy reform and cognitive radio, Proc. IEEE 97 (4) (2009) 708–719.

[5] Y. Saleem, M.H. Rehmani, Primary radio user activity models for cognitive radio networks: A survey, J. Netw. Comput. Appl. 43 (2014) 1–16.

[6] B. Wang, et al., Game theory for cognitive radio networks: An overview, Comput. Netw. 54 (14) (2010) 2537–2561.

[7] A. Ahmad, et al., A survey on radio resource allocation in cognitive radio sensor networks, Commun. Surv. Tutor. 17 (2) (2015) 888–917.

[8] C.H. de Lima, et al., Coordination mechanisms for self-organizing femtocells in two-tier coexistence scenarios, IEEE Trans. Wireless Commun. 11 (6) (2012) 2212–2223.

[9] L.D. Xu, et al., Internet of things in industries: A survey, IEEE Trans. Ind. Inf. 10 (4) (2014) 2233–2243.

[10] S.D.T. Kelly, et al., Towards the implementation of IoT for environmental condition monitoring in homes, IEEE Sens. J. 13 (10) (2013) 3846–3853.

[11] A. Bradai, et al., EMCOS: Energy-efficient mechanism for multimedia streaming over cognitive radio sensor networks, Pervasive Mob. Comput. 22 (2015) 16–32.

[12] S.F. Bush, Smart Grid: Communication-enabled Intelligence for the Electric Power grid, John Wiley & Sons, 2014.

[13] P.H.J. Nardelli, et al., Models for the modern power grid, EPJ-ST 223 (12) (2014) 2423–2437.

[14] Y. Yan, et al., A survey on smart grid communication infrastructures: Motivations, requirements and challenges, Commun. Surv. Tutor. 15 (1) (2013) 5–20.

[15] M. Kuzlu, et al., Communication network requirements for major smart grid applications in HAN, NAN and WAN, Comput. Netw. 67 (2014) 74–88.

[16] R.C. Qiu, et al., Cognitive radio network for the smart grid: experimental system architecture, control algorithms, security, and microgrid testbed, IEEE Trans. Smart Grid 2 (4) (2011) 724–740.

[17] G. Shah, et al., A cross-layer QoS-aware communication framework in cognitive radio sensor networks for smart grid applications, IEEE Trans. Ind. Inf. 9 (3) (2013) 1477–1485.

[18] A.A. Khan, et al., Cognitive radio for smart grids: Survey of architectures, spectrum sensing mechanisms, and networking protocols, Commun. Surv. Tutor. (2015).

[19] R. Yu, et al., Cognitive radio based hierarchical communications infrastructure for smart grid, IEEE Netw. 25 (5) (2011) 6–14.

[20] V.C. Gungor, D. Şahin, Cognitive radio networks for smart grid applications: A promising technology to overcome spectrum inefficiency, IEEE Veh. Technol. Mag. 7 (2) (2012) 41–46.

[21] J.J. Nielsen, et al., What can wireless cellular technologies do about the upcoming smart metering traffic? IEEE Commun. Mag. 53 (9) (2015) 41–47.

[22] Y. Zhang, et al., Cognitive machine-to-machine communications: visions and potentials for the smart grid, IEEE Netw. 26 (3) (2012) 6–13.

[23] J. Wildman, et al., On the joint impact of beamwidth and orientation error on throughput in directional wireless Poisson networks, IEEE Trans. Wirel. Commun 13 (12) (2014) 7072–7085.

[24] J. Andrews, Interference cancellation for cellular systems: A contemporary overview, IEEE Wirel. Commun. 12 (2) (2005) 19–29.

[25] M. Haenggi, Stochastic Geometry for Wireless Networks, Cambridge University Press, 2012.

[26] P.H.J. Nardelli, et al., Optimal transmission capacity of ad hoc networks with packet retransmissions, IEEE Trans. Wirel. Commun. 11 (8) (2012) 2760–2766.

[27] P.H.J. Nardelli, Throughput optimization in wireless networks under stability and packet loss constraints, IEEE Trans. Mob. Comput. 13 (8) (2014) 1883–1895.

[28] P.H.J. Nardelli, et al., Throughput analysis of cognitive wireless networks with Poisson distributed nodes based on location information, Ad Hoc Netw. (2015).

[29] J.Z. Kolter, M.J. Johnson, REDD: a public data set for energy disaggregation research, Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA 25 (Citeseer) (2011) 59–62.

[30] J.Z. Kolter, [Online]. Available: http://redd.csail.mit.edu/ (accessed 15.03.15).

[31] P. Cardieri, Modeling interference in wireless ad hoc networks, Commun. Surv. Tutor. 12 (4) (2010) 551–572.

[32] S. Weber, An overview of the transmission capacity of wireless networks, IEEE Trans. Commun. 58 (12) (2010) 3593–3604.

[33] F. Baccelli, et al., Interference networks with point-to-point codes, IEEE Trans. Inf. Theory 57 (5) (2011) 2582–2596.

[34] S. Weber, et al., The effect of fading, channel inversion, and threshold scheduling on ad hoc networks, IEEE Trans. Inf. Theory 53 (11) (2007) 4127–4149.

[35] [Online]. Available: http://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.optimize.fsolve.html (accessed 15.03.15).

[36] A. Goldsmith, Wireless Communications, Cambridge University Press, 2005.

[37] M. Pipattanasomporn, et al., Load profiles of selected major household appliances and their demand response opportunities, IEEE Trans. Smart Grid 5 (2) (2014) 742–750.

[38] E.-K. Lee, et al., Physical layer security in wireless smart grid, IEEE Commun. Mag. 50 (8) (2012) 46–52.

[39] G. Strbac, Demand side management: Benefits and challenges, Energy policy 36 (12) (2008) 4419–4426.

[40] Y. Tsividis, Event-driven data acquisition and digital signal processing: A tutorial, IEEE Trans. Circ. Syst. II, Exp. Br. 57 (8) (2010) 577–581.

[41] M. Simonov, et al., Gathering process data in low-voltage systems by enhanced event-driven metering, IEEE Syst. J PP (99) (2015) 1–12.

[42] J. Munkhammar, et al., Characterizing probability density distributions for household electricity load profiles from high-resolution electricity use data, Appl. Energ. 135 (2014) 382–390.

[43] I. Sajjad, et al., Probabilistic generation of time-coupled aggregate residential demand patterns, IET Gener. Transm. Distrib. 9 (9) (2015) 789–797.

[44] M.A. Alsheikh, et al., Machine learning in wireless sensor networks: Algorithms, strategies, and applications, Commun. Surv. Tutor. 16 (4) (2014) 1996–2018.

[45] A. Aijaz, H. Aghvami, PRMA based cognitive machine-to-machine communication in smart grid networks, IEEE Trans. Veh. Technol PP (99) (2014). 1–1.

[46] A. Aijaz, et al., CORPL: A routing protocol for cognitive radio enabled AMI networks, IEEE Trans. Smart Grid 6 (1) (2015) 477–485.

[47] H. Alves, et al., Enhanced transmit antenna selection scheme for secure throughput maximization without CSI at the transmitter and its applications on smart grids, IEEE Trans. Inf. Forensics Secur. (2015).[Online]. Available: http://arxiv.org/pdf/1511.06518

Pedro Henrique Juliano Nardelli received the B.S. and M.Sc. degrees in electrical engineering from the State University of Campinas, Brazil, in 2006 and 2008, respectively. In 2013 he received his doctoral degree from University of Oulu, Finland, and State University of Campinas following a dual-degree agreement. Nowadays he holds a postdoctoral position at University of Oulu, and his studies are mainly focused on the efficiency of wireless networks and spatio-temporal dynamics of complex systems.

Mauricio de Castro Tomé received the B.S and M.Sc. in electrical engineering from the State University of Campinas, Brazil, in 2011 and 2014, respectively. He is currently pursuing his Ph.D. degree at the University of Oulu, Finland, researching power and energy measurements and load estimation in smart grids.

Hirley Alves received the B.Sc., M.Sc. and D.Sc. degrees from Federal University of Technology - Paraná (UTFPR), Brazil, in 2010, 2011 and 2015, respectively. Hirley has jointly graduated from University of Oulu, and received his D.Sc. in 2015. Hirley is postdoctoral researcher at Centre for Wireless Communications (CWC), Oulu. His current research focuses on: performance analysis of full-duplex networks and relaying and its applications on 5G, and smart grids.

Carlos H. M. de Lima received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Ceará, Fortaleza, Brazil, in 2002 and 2004, respectively. In 2013 he is received his D.Sc. degree in of Telecommunications Engineering from University of Oulu, Finland. He is currently an assistant professor at Sao Paulo State University, São João da Boa Vista-SP, Brazil, and a member of the research staff with the Centre for Wireless Communications, University of Oulu. From 2000 to 2005, he was a Research Scientist with the Wireless Telecommunications Research Group (GTEL), Fortaleza. In 2005, he was a Visiting Researcher with the Ericsson Research Center, Lulea, Sweden, engaged in power control techniques for enhanced high-speed packet access systems. In 2006, he was with Nokia Institute of Technology (INdT), Brazil, engaged in the evaluation of the system performance of WiMAX systems. His research interests include statistical signal processing and analysis of interference networks using stochastic geometry.

Matti Latva-aho was born in Kuivaniemi, Finland in 1968. He received the M.Sc., Lic.Tech. and Dr. Tech (Hons.) degrees in Electrical Engineering from the University of Oulu, Finland in 1992, 1996 and 1998, respectively. From 1992 to 1993, he was a Research Engineer at Nokia Mobile Phones, Oulu, Finland. During the years 1994–1998 he was a Research Scientist at Telecommunication Laboratory and Centre for Wireless Communications at the University of Oulu. Currently he is the Department Chair Professor of Digital Transmission Techniques and Head of Department at the University of Oulu, Department for Communications Engineering. Prof. Latva-aho was Director of Centre for Wireless Communications at the University of Oulu during the years 1998-2006. His research interests are related to mobile broadband wireless communication systems. Prof. Latva-aho has published over 200 conference or journal papers in the field of wireless communications. He has been TPC Chairman for PIMRC06, TPC Co-Chairman for ChinaCom07 and General Chairman for WPMC08. He acted as the Chairman and vice-chairman of IEEE Communications Finland Chapter in 2000–2003.