SWIPT Signalling over Complex AWGN Channels with Two Nonlinear Energy Harvester Models

Morteza Varasteh, Borzoo Rassouli, Hamdi Joudeh and Bruno Clerckx

Department of Electrical and Electronic Engineering, Imperial College London, London, U.K.

{m.varasteh12; b.rassouli12; hamdi.joudeh10; b.clerckx}@imperial.ac.uk.

Abstract—Simultaneous Wireless Information and Power Transfer (SWIPT) is subject to nonlinearity at the energy harvester that leads to significant changes to transmit signal designs compared to conventional wireless communications. In this paper, the capacity of a discrete time, memoryless and complex Additive White Gaussian Noise (AWGN) channel in the presence of a nonlinear energy harvester at the receiver is studied. Considering the two common nonlinear energy harvester models introduced in the literature, two sets of constraints are considered. First the capacity is studied under average power (AP), peak amplitude (PA) and receiver delivery power (RDP) constraints. The RDP constraint is modelled as a linear combination of even-moment statistics of the channel input being larger than a threshold. It is shown that the capacity of an AWGN channel under AP and RDP constraints is the same as the capacity of an AWGN channel under an AP constraint, however, depending on the two constraints, it can be either achieved or arbitrarily approached. It is also shown that under AP, PA and RDP constraints, the amplitude of the optimal inputs is discrete with a finite number of mass points. Next, the capacity is studied under AP, PA and output outage probability (OOP) constraints. OOP is modelled as satisfying a certain probability inequality for the amplitude of the received signal being outside of a given interval. Similarly, it is shown that the amplitude of the optimal input is discrete with a finite number of mass points.

I. INTRODUCTION

One of the major efforts in a Simultaneous Wireless Information and Power Transfer (SWIPT) architecture is to increase the Direct-Current (DC) power and information rate at the output of the energy harvester and information decoder, respectively, without increasing the transmit power. As one of the primary works in the information theory literature, Varshney studied this problem in [1], in which he characterized the capacity-power function for a point-to-point discrete memoryless channel. The harvester, known as rectenna, is composed of an antenna followed by a rectifier.[1] In [2], [3], it is shown that as a consequence of the rectifier nonlinearity, the RF-to-DC conversion efficiency is a function of rectenna's structure, as well as its input waveform. Accordingly, in order to maximize rectenna's DC power output, a systematic waveform design is crucial to make the best use of an available RF spectrum [3]. Following [3], it was shown in [4], [5] that the choice of a suitable input distribution (and therefore modulation and waveform) for SWIPT is affected by the rectifier nonlinearity

[1]In the literature, the rectifier is usually considered as a nonlinear device (usually a diode) followed by a low-pass filter. The diode is the main source of nonlinearity induced in the system.

and motivates the study of the capacity of AWGN channels under nonlinear power constraints.

In the literature, so far, depending on the application and available resources, two different models of the energy harvester nonlinearity are used. The first model is based on Taylor expansion of the diode characteristic function and is introduced in [3]. It is shown that the harvested power is a function of summation of even moments of the received RF signal, which can be approximated with an acceptable level of accuracy by truncating it to the second and fourth moments. This model is appropriate for practical low power applications (with input power to the energy harvester of the order of 0dBm or less). The other model, which is a function of the RF power of the received signal, is introduced in [6]. In this model, the rectenna's input/output power relationship is expressed in terms of a logistic (sigmoidal) function. This model is appropriate in applications where there is little guarantee for the energy harvester to operate below the diode breakdown edge.

The capacity of deterministic, complex and real, discrete-time memoryless AWGN channels has been investigated in the literature under various constraints, extensively. It seems that the linear AWGN channel subject to transmit average power constraint is an exception where the optimal input is Gaussian distributed [7], and under many other constraints, the optimal input leads to discrete inputs. To mention a few, [8] for a real AWGN channel with average power and amplitude constrained inputs, [9] for complex AWGN channels with average and peak-power constraint and [10] for complex Rayleigh-fading channel when no channel state information (CSI) is assumed either at the receiver or the transmitter.

Motivated by the two nonlinear models in the literature, we provide a step closer at identifying the fundamental limits of SWIPT structures. In this paper, we study a complex and discrete time memoryless AWGN channel under two sets of constraints. First, in order to tackle the nonlinear model introduced in [3], we study the capacity problem under average power (AP), peak amplitude (PA) and receiver delivery power (RDP) constraints. Next, to tackle the model in [6], we study the capacity problem under average power (AP), peak amplitude (PA) and output outage probability (OOP) constraints. First, focusing on the nonlinear model of the harvester in [3], we show that the capacity of an AWGN channel under AP and RDP constraints is the same as the capacity of an AWGN channel under AP constraint. However, depending on the two constraints, the capacity can be either achieved or

approached arbitrarily (irrespectively of the RDP constraint). We show that in line with the results reported in [8]–[11] the optimal input distributions are discrete with a finite number of mass points. The system model studied in this paper focuses on the nonlinearities at the receiver (over complex AWGN channels) and indeed can be considered as a reciprocal of [11], where the main focus is on the nonlinearities at the transmitter (transmit nonlinear constraints as well as nonlinear channel inputs over real AWGN channels). Second, focusing on the nonlinear model of the harvester in [6], we study the capacity problem under AP, PA and OOP constraint, where similar results (discrete inputs with a finite number of mass points) are obtained.

Organization: In Section II, we introduce the system model and define the channel capacity problem studied here. In Section III, we introduce the main results of the paper. In Section IV, numerical results are illustrated. The proofs of the provided results are summarized due to space limit, and where it is necessary, we refer to the details in the extended version of the paper [12].

Notations: Throughout this paper, the standard circular symmetric complex Gaussian (CSCG) distribution is denoted by $\mathcal{CN}(0,1)$. $F_{\boldsymbol{r}}(r)$ and $f_{\boldsymbol{r}}(r)$ denote, respectively, cumulative distribution function (CDF) and the probability density function (pdf) of thr random variable $\boldsymbol{r}$, with its support denoted as $\mathrm{supp}\{\boldsymbol{r}\}$. We define the kernel $K(R,r) \triangleq Re^{-\frac{R^2+r^2}{2}} I_0(rR)$, where $I_0(x) = 1/\pi \int_0^\pi e^{x\cos(\theta)} d\theta$ is the modified Bessel function of the first kind and order zero. The marcum Q-function is defined as $Q(r,A) \triangleq \int_A^\infty K(R,r) dR$. The error function is defined as $\mathrm{erf}(x) = 2/\sqrt{\pi} \int_0^x e^{-t^2} dt$.

II. System Model, Problem Definition and Preliminaries

Consider the following complex representation of a discrete-time AWGN channel,

$$\boldsymbol{y}_k = \boldsymbol{x}_k + \boldsymbol{n}_k, \tag{1}$$

where $\{\boldsymbol{y}_k\}$, $\{\boldsymbol{x}_k\}$ and $\{\boldsymbol{n}_k\}$ represent the sequences of complex-valued samples of the channel output, input and AWGN, respectively, and k is the discrete-time index. The real and imaginary parts of the signal $\{\boldsymbol{y}_k\}$ indicate the Inphase and Quadrature components, respectively. The noise samples $\{\boldsymbol{n}_k\}$ are assumed to be CSCG distributed as $\mathcal{CN}(0,2)$.[2]

A. Capacity under AP, PA and RDP constraints

The nonlinear model for the rectenna introduced in [3] is a function of even moments of the received Radio Frequency (RF) signal. In [12], it is shown that the modelled harvested power can be lower bounded by considering the even moments of the baseband equivalent of the RF signal. Motivated by that,

the capacity of a discrete-time complex AWGN channel [13, Ch. 7] under AP, PA and RDP is given by

$$C_{P_a, P_d, r_p} \triangleq \sup_{p_{\boldsymbol{x}}(x)} \quad I(\boldsymbol{x}; \boldsymbol{y})$$
$$\text{s.t.} : \quad \begin{cases} \mathbb{E}[|\boldsymbol{x}|^2] \leq P_a, \\ \mathbb{E}[g(|\boldsymbol{x}|)] \geq P_d, \\ |\boldsymbol{x}| \leq r_p, \end{cases} \tag{2}$$

where P_a, P_d and r_p are the transmitter AP, PA and RDP constraints, respectively. $g(\cdot)$ is assumed to be a continuous positive function having the form of

$$g(r) = \sum_{i=0}^n \alpha_i r^{2i}, \ r \geq 0, \tag{3}$$

where $n \geq 2$ is an arbitrary integer. Note that since $g(r)$ is assumed to be a positive function, we have $\alpha_n > 0$, and hence, $\lim_{r\to\infty} g(r) = \infty$.[3]

B. Capacity under AP, PA and OOP constraints

The model for energy harvester proposed in [6], captures both the threshold effect[4] (the power level below which the rectifier turns off and denoted by A_l^2) and the saturation effect (the power level above which the rectifier's power gain remains constant and denoted by A_u^2). From an energy harvesting point of view, it is most favourable to keep the energy harvester operating without getting saturated. Motivated by this nonlinear model, we introduce the OOP constraint as $\Pr(|\boldsymbol{x}| \geq A_u \cup |\boldsymbol{x}| \leq A_l) \leq 1-\epsilon$ for $0 \leq A_l < A_u \leq \infty$ and $0 \leq \epsilon < 1$. We note that since this criterion is equivalent to $\Pr(A_l \leq |\boldsymbol{y}| \leq A_u) \geq \epsilon$, in the following for simplicity we consider the latter as the OOP constraint. The capacity of a complex discrete-time AWGN channel [13, Chapter 7] under AP, PA and OOP is given by

$$C_{P_a, A_l, A_u, \epsilon, r_p} \triangleq \sup \quad I(\boldsymbol{x}; \boldsymbol{y})$$
$$p_{\boldsymbol{x}}(x) : \quad \begin{cases} \mathbb{E}[|\boldsymbol{x}|^2] \leq P_a, \\ |\boldsymbol{x}| < r_p, \\ \Pr(A_l \leq |\boldsymbol{y}| \leq A_u) \geq \epsilon, \end{cases} \tag{4}$$

C. Preliminaries

Expressing $I(\boldsymbol{x}; \boldsymbol{y})$ in terms of differential entropies, we have

$$I(\boldsymbol{x}; \boldsymbol{y}) = h(\boldsymbol{y}) - \log 2\pi e. \tag{5}$$

Therefore, the problems in (2) and (4) are equivalent to the supremization of differential entropy $h(\boldsymbol{y})$ subject to the corresponding constraints in (2) and (4), respectively. Using polar coordinates[5] $\boldsymbol{x} = \boldsymbol{r} e^{i\boldsymbol{\theta}}$ and $\boldsymbol{y} = \boldsymbol{R} e^{i\boldsymbol{\phi}}$ ($\boldsymbol{r}, \boldsymbol{R} \geq 0$ and $\boldsymbol{\theta}, \boldsymbol{\phi} \in [0, 2\pi)$) and following the same steps in [9], we have

$$h(\boldsymbol{y}) \leq - \int_0^\infty f_{\boldsymbol{R}}(R; F_{\boldsymbol{r}}) \ln \frac{f_{\boldsymbol{R}}(R; F_{\boldsymbol{r}})}{R} dR + \ln 2\pi, \tag{6}$$

where $f_{\boldsymbol{R}}(R; F_{\boldsymbol{r}})$ is the pdf of $\boldsymbol{R}$ induced by $F_{\boldsymbol{r}}$ given as $f_{\boldsymbol{R}}(R; F_{\boldsymbol{r}}) = \int_0^{r_p} K(R, r) dF_{\boldsymbol{r}}(r)$. Note that by selecting $\boldsymbol{r}$ and $\boldsymbol{\theta}$ independent with uniformly distributed $\boldsymbol{\theta}$ over $[0, 2\pi)$, the upperbound in (6) becomes tight (for more details see [9]) and we have $f_{\boldsymbol{R},\boldsymbol{\phi}}(R, \phi) = \frac{1}{2\pi} f_{\boldsymbol{R}}(R; F_{\boldsymbol{r}})$.

In (4), the OOP constraint can be equivalently written as

$$\Pr(A_l \le |\boldsymbol{y}| \le A_u) = \int\limits_0^\infty (Q(r, A_l) - Q(r, A_u))\, dF_{\boldsymbol{r}}(r). \quad (7)$$

Therefore, the optimization problem in (2) and (4) are reduced to the following problems, respectively

$$C_{P_a, P_d, r_p} = \sup_{F_{\boldsymbol{r}} \in \Omega_1 \cap \Omega_2} H(F_{\boldsymbol{r}}), \quad (8)$$

$$C_{P_a, A_l, A_u, \epsilon, r_p} = \sup_{F_{\boldsymbol{r}} \in \Omega_1 \cap \Omega_3} H(F_{\boldsymbol{r}}), \quad (9)$$

where $F_{\boldsymbol{r}}$ is such that $F_{\boldsymbol{r}}(0^-) = 0, F_{\boldsymbol{r}}(r_p) = 1$ and $H(F_{\boldsymbol{r}})$, Ω_1, Ω_2 and Ω_3 are given as

$$H(F_{\boldsymbol{r}}) \triangleq -\int\limits_0^\infty f_{\boldsymbol{R}}(R; F_{\boldsymbol{r}}) \ln \frac{f_{\boldsymbol{R}}(R; F_{\boldsymbol{r}})}{R} dR, \quad (10)$$

$$\Omega_1 = \left\{ F_{\boldsymbol{r}} : \int_0^{r_p} r^2 dF_{\boldsymbol{r}}(r) \le P_a \right\}, \quad (11a)$$

$$\Omega_2 = \left\{ F_{\boldsymbol{r}} : \int_0^{r_p} g(r) dF_{\boldsymbol{r}}(r) \ge P_d \right\}, \quad (11b)$$

$$\Omega_3 = \left\{ F_{\boldsymbol{r}} : \int_0^{r_p} (Q(r, A_l) - Q(r, A_u)) dF_{\boldsymbol{r}}(r) \ge \epsilon \right\}. \quad (11c)$$

III. MAIN RESULTS

In this section, we provide the main results of this paper regarding the problems introduced in (8) and (9).

A. Capacity under AP, PA and RDP constraints

In the following, we first characterize the capacity in (8) when the channel input amplitude constraint is $r_p = \infty$. In the next theorem, we study the capacity problem in (8) when $r_p < \infty$. We accordingly, derive the necessary and sufficient condition for the optimal distributions achieving the capacity.

Theorem 1. *The capacity of the channel in (1) under AP, RDP constraints and $r_p = \infty$, i.e., $C_{P_a, P_d, \infty}$ is characterized as*

$$C_{P_a, P_d, \infty} = \log(1 + P_a/2). \quad (12)$$

If $P_d \le P_R$, the capacity $C_{P_a, P_d, \infty}$ is achieved by a unique input distributed as $\boldsymbol{x} \sim \mathcal{CN}(0, P_a)$, and for $P_d > P_R$, the capacity $C_{P_a, P_d, \infty}$ is not achieved, however, can be approached arbitrarily, where $P_R = \frac{1}{P_a} \int_0^\infty r g(r) e^{-\frac{r^2}{2P_a}} dr$ is the RDP corresponding to CSCG input.

Proof: It can be verified that the space $\Omega_1 \cap \Omega_2$ is convex. For $r_p = \infty$, the space $\Omega_1 \cap \Omega_2$ is not compact (see [12, Lem. 3] for a counterexample contradicting the compactness). It is shown in [12, Lem. 10] that for every $F_{\boldsymbol{r}} \in \Omega_1 \cap \Omega_2$, $H(F_{\boldsymbol{r}})$ exists, and is continuous, concave and weakly differentiable. Strict concavity of $H(F_{\boldsymbol{r}})$ over $\Omega_1 \cap \Omega_2$ is proved by noting that the integral transform $f_{\boldsymbol{R}}(R; F_{\boldsymbol{r}}) = \int_0^\infty K(R, r) dF_{\boldsymbol{r}}(r)$ is invertible (for the proof see [9, App. II]). For $P_d \le P_R$, the RDP

constraint is inactive and the optimal input is $\boldsymbol{x} \sim \mathcal{CN}(0, P_a)$. For $P_d > P_R$, it is easily verified that for a fixed P_a, the capacity $C_{P_a, P_d, \infty}$ is non-increasing with P_d. Therefore, we have $C_{P_a, P_R, \infty} \ge C_{P_a, P_d, \infty}$ for $P_d > P_R$. Also due to the strict concavity of $H(F_{\boldsymbol{r}})$, the capacity achieving distributions are unique. Accordingly, we have $C_{P_a, P_R, \infty} > C_{P_a, P_d, \infty}$. In the following, we show that there exist distributions that can approach the $C_{P_a, P_R, \infty}$ arbitrarily.

Consider the following sequence of distribution functions

$$F_{\boldsymbol{r}_l}(r) = \begin{cases} 0 & r < 0 \\ 1 - \frac{1}{l^2} & 0 \le r < \sqrt{P_a}l \\ 1 & r \ge \sqrt{P_a}l \end{cases}, \; l = 2, \dots. \quad (13)$$

Note that $\mathbb{E}_{F_{\boldsymbol{r}_l}}[\boldsymbol{r}_l^2] = P_a$, hence, satisfying the AP constraint. Also, for the RDP constraint we have

$$P_{d,l} \triangleq \mathbb{E}_{F_{\boldsymbol{r}_l}}[g(\boldsymbol{r}_l)] = \alpha_0 + \alpha_1 P_a + \sum_{i=2}^n \alpha_i P_a^i l^{2i-2}. \quad (14)$$

Since $n \ge 2$ by construction, it is guaranteed that there exists an integer number L, such that for $l > L$, $P_{d,l} \ge P_d$ (note that $P_{d,l}$ increases with l). Due to convexity of the optimization space, time sharing is valid in our system model. Hence, we can construct a complex input $\boldsymbol{x}_l$ with its phase $\boldsymbol{\theta}_l$ uniformly distributed over $[0, 2\pi)$ and its amplitude $\boldsymbol{r}_l$ distributed according to the following CDF

$$F_{\boldsymbol{r}_{ts}}(r) = (1 - \tau) F_{\boldsymbol{r}_R}(r) + \tau F_{\boldsymbol{r}_l}(r), \tau \in (0, 1), l > L, \quad (15)$$

where the subscript ts in $F_{\boldsymbol{r}_{ts}}$ stands for time-sharing and $F_{\boldsymbol{r}_R}(r)$ corresponds to the CDF of the Rayleigh distribution (recall that amplitude of a CSCG distributed complex random variable has a Rayleigh CDF). By choosing $\tau = (P_d - P_R)/(P_{d,l} - P_R)$, it is easy to verify that $0 < \tau < 1$ and the constraints

$$\mathbb{E}_{F_{\boldsymbol{r}_{ts}}}[\boldsymbol{r}_{ts}^2] \le P_a, \; \mathbb{E}_{F_{\boldsymbol{r}_{ts}}}[g(\boldsymbol{r}_{ts})] \ge P_d, \quad (16)$$

are both satisfied. On the other hand, due to strict concavity of the entropy $H(F_{\boldsymbol{r}})$, we have

$$H(F_{\boldsymbol{r}_{ts}}) > (1 - \tau) H(F_{\boldsymbol{r}_R}) + \tau H(F_{\boldsymbol{r}_l}), \tau \in (0, 1), l > L. \quad (17)$$

For a given P_d, we can increase l arbitrarily. Since $P_{d,l}$ increases with l, therefore τ can be made arbitrarily close to zero. Rewriting (17), we have

$$H(F_{\boldsymbol{r}_R}) > H(F_{\boldsymbol{r}_{ts}}) > (1 - \tau) H(F_{\boldsymbol{r}_R}) + \tau H(F_{\boldsymbol{r}_l}), \quad (18)$$

where by letting τ tend to zero (equivalently letting $P_{d,l} \to \infty$) the result of Theorem 1 is concluded. $\square$

From the result of Theorem 1, it is verified that for $n \ge 2$ in (3), the capacity of an AWGN channel in (1) for $r_p = \infty$ is independent of the value of the RDP constraint, i.e., P_d. That is, given P_a, the capacity $C_{P_a, P_d, \infty}$ is unchanged with P_d.

Theorem 2. *The optimal distribution denoted by $F_{\boldsymbol{r}_1^o}$ achieving the capacity C_{P_a, P_d, r_p} for $r_p < \infty$, is unique and the corresponding set of points of increase is finite (,i.e., the cardinality of the random variable $\boldsymbol{r}_1^o$ is finite). Furthermore,*

$F_{\boldsymbol{r}_1^o}$ is optimal if and only if there are unique parameters $\mu_1, \mu_2 \geq 0$ for which

$$h(r; F_{\boldsymbol{r}_1^o}) - \mu_1 r^2 + \mu_2 g(r) - K_1 = 0, \forall r \in \mathrm{supp}\{\boldsymbol{r}_1^o\}, \quad (19a)$$

$$h(r; F_{\boldsymbol{r}_1^o}) - \mu_1 r^2 + \mu_2 g(r) - K_1 \leq 0, \forall r \in [0, r_p], \quad (19b)$$

where $K_1 \triangleq H(F_{\boldsymbol{r}_1^o}) - \mu_1 P_a + \mu_2 P_d$ and

$$h(r; F_{\boldsymbol{r}}) = -\int_0^\infty K(R, r) \log \frac{f(R, F_{\boldsymbol{r}})}{R} dR. \quad (20)$$

Proof: For $r_p < \infty$, the probability optimization space $\Omega_1 \cap \Omega_2$ is convex and compact in the Levi metric [8]. This along with the existence, continuity, strict concavity and weak differentiability of $H(F_{\boldsymbol{r}})$ over $F_{\boldsymbol{r}} \in \Omega_1 \cap \Omega_2$ (see [12, Lem. 10]) qualifies applying the Lagrangian theorem [14, Sec. 7.4]. Hence, the necessary and sufficient condition for an input $\boldsymbol{r}_1^0$ to be optimal is

$$\int_0^{r_p}(h(r; F_{\boldsymbol{r}_1^o}) - \mu_1 r^2 + \mu_2 g(r)) dF_{\boldsymbol{r}}(r) \leq K_1, \forall F_{\boldsymbol{r}} \in \Omega_1 \cap \Omega_2. \quad (21)$$

Noting that $K(R, r)$ is an analytic function (with respect to both of its arguments) and following a similar approach as in [8], it can be verified that the amplitude of the optimal input $\boldsymbol{r}_1^0$ is discrete with a finite number of mass points (see [12] for more details). Note that for the optimal complex input $\boldsymbol{x}$, it is enough to consider the phase $\boldsymbol{\theta}$ uniformly distributed over $[0, 2\pi)$. $\square$

Remark 1. *By expanding $h(r; F_{\boldsymbol{r}})$ in (20) (see [12, App. C] for more details), we have*

$$\int_0^\infty K(R, r) \log \frac{R}{f(R, F_{\boldsymbol{r}})} dR > -2.$$

Rewriting the KKT condition, for the inequality condition in (19) we get

$$0 \leq \mu_2 \leq \frac{K_1 + 2 + \mu_1 r^2}{g(r)}, \ r \in [0, r_p]. \quad (22)$$

Since by definition, the function $g(r)$ grows faster than r^2, and (22) is valid for any $r \in [0, r_p]$, by substituting $r = r_p$, we have $\mu_2 \to 0$ as $r_p \to \infty$. The intuition behind this is as follows. μ_2 can be considered as the opposite sign of $\partial C_{P_a, P_d, r_p}/\partial r_p$. As r_p increases, C_{P_a, P_d, r_p} approaches $C_{P_a, P_d, \infty}$. From Theorem 1, we already know that capacity $C_{P_a, P_d, \infty}$ is unchanged for any $P_d < \infty$. Therefore, $\partial C_{P_a, P_d, r_p}/\partial r_p \to 0$ as $r_p \to \infty$.

B. Capacity under AP, PA and OOP constraints

In the following, we derive the necessary and sufficient condition for the optimal distributions achieving the capacity in (9), when the input amplitude constraint is $r_p < \infty$.

Theorem 3. *The optimal distribution denoted by $F_{\boldsymbol{r}_2^o}$ achieving the capacity $C_{P_a, A_l, A_u, \epsilon, r_p}$ for $r_p < \infty$, is unique and the corresponding set of points of increase is finite (,i.e., the cardinality of the random variable $\boldsymbol{r}_2^o$ is finite). Furthermore,*

$F_{\boldsymbol{r}_2^o}$ is optimal if and only if there are unique parameters $\lambda_1, \lambda_2 \geq 0$ for which

$$h(r; F_{\boldsymbol{r}_2^o}) - \lambda_1 r^2 + \lambda_2 (Q(r, A_l) - Q(r, A_u)) = K_2, \forall r \in \mathrm{supp}\{\boldsymbol{r}_2^o\}, \quad (23a)$$

$$h(r; F_{\boldsymbol{r}_2^o}) - \lambda_1 r^2 + \lambda_2 (Q(r, A_l) - Q(r, A_u)) \leq K_2, \forall r \in [0, r_p], \quad (23b)$$

where $K_2 \triangleq H(F_{\boldsymbol{r}_2^o}) - \lambda_1 P_a + \lambda_2 \epsilon$.

Proof: Noting that $Q(r, A_l) - Q(r, A_u)$ is a bounded and continuous function of r, the convexity and compactness of the optimization probability space $\Omega_1 \cap \Omega_3$ is verified in a similar approach as in [10, Appendix A][6]. Therefore, by applying the Lagrangian theorem and following the same argument as in [8], it is shown that it is necessary and sufficient for the optimal input $\boldsymbol{r}_2^o$ to satisfy (23). Due to the fact that the function $Q(z, A_l) - Q(z, A_u)$ is analytic over z, using identity theorem and the fact that the distribution of amplitude of the received signal is continuous for every $F_{\boldsymbol{r}} \in \Omega_1 \cap \Omega_3$, it is shown that the optimal input does not contain a limit point in its support. This along with the fact that the support of the input is finite $r_p < \infty$, yields the desired result. $\square$

Note that the results in (19) and (23) are important in the sense that they can be utilized to obtain the optimal distributions using numerical programming.

IV. NUMERICAL RESULTS

In this section, we provide some numerical illustrations of the results obtained in Section III. In the following, we first summarize the steps in obtaining the optimal inputs, and next, we illustrate the obtained numerical results.

To obtain the optimal inputs corresponding to the problems in (2) and (4), we resort to numerical programming. Accordingly, we solve the optimization problems using the interior-point algorithm implemented by the fmincon function in MATLAB software. Note that, since we already know that the optimal distribution is discrete with a finite number of mass points, the numerical optimization is over the position, the probabilities and the number of the mass points. Hence, there are $2m$ parameters to be optimized, where m denotes the number of the mass points. Given the constraints, the algorithm is fed with an input of two mass points, i.e., $m = 2$ and a random guess for the respective probabilities and positions of the mass points. Once the algorithm outputs the local optimal solution for the positions and their respective probabilities,[7] the answer is validated by checking the constraints and the corresponding necessary and sufficient KKT conditions in (19) or (23). If the conditions are not satisfied, the initial guess is changed. We continue changing the initial guess for a large number of times. If the KKT conditions are not satisfied yet, the number of mass points is increased by one. The algorithm runs until such an input is found.

---

[6]We note that the convexity and compactness hold valid for $r_p \leq \infty$

[7]Note that despite the fact that the problem is concave with respect to probability laws, however, for a given number of mass points m, the problem is not concave and the obtained solution is not guaranteed to be a global one.
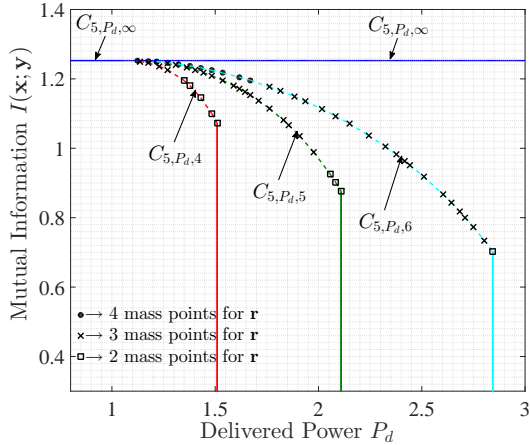
Figure 1. Mutual information $I(\boldsymbol{x}; \boldsymbol{y})$ corresponding to the optimal solutions of (2) with respect to different values of the RDP constraint P_d with PA constraints $r_p = 4$, 5, 6 and $r_p = \infty$ and AP constraint $P_a = 5$.

Figure 2. Mutual information $I(\boldsymbol{x}; \boldsymbol{y})$ corresponding to the optimal solutions of (4) with respect to different values of the OOP constraint ϵ with PA constraint $r_p = 6$, AP constraint $P_a = 10$.

Illustration of the numerical results: In Fig. 1, simulation results for the problem (2) with $P_a = 5$ and $g(r) = 0.01(r^4 + r^2 + 1)$ are illustrated. The horizontal solid line related to $C_{5,P_d,\infty}$ corresponds to the AWGN channel capacity under an AP constraint $P_a = 5$ achieved by only a CSCG distribution. The horizontal dashed line related to $C_{5,P_d,\infty}$ corresponds to the capacity under an AP constraint $P_a = 5$, which is not achievable, however, can be approached arbitrarily (see Theorem 1). $C_{5,P_d,4}$, $C_{5,P_d,5}$ and $C_{5,P_d,6}$ correspond to the optimal solution in (2) for $r_p = 4$, 5 and 6, respectively. It is observed that by increasing the PA constraint r_p, the capacity tends to the one corresponding to $r_p = \infty$. This observation is inline with Remark 1, that increasing r_p, reduces the dependency of the capacity on r_p. Note that given the value of r_p, the amount of maximum RDP at the receiver is limited. This is the reason for the vertical lines corresponding to $C_{5,P_d,4}$, $C_{5,P_d,5}$ and $C_{5,P_d,6}$. It is observed that the number of optimal mass points m for $\boldsymbol{r}$ is decreased with P_d.

In Fig. 2, similar results to the problem (4) are obtained with AP $P_a = 10$, PA $r_p = 6$ for different values of A_l, A_u and ϵ. It is observed that by increasing ϵ (equivalently decreasing the outage probability), the rate is decreased. It is also observed that the number of mass points decreases with ϵ.

We note that, the algorithm used for finding the optimal distributions is extremely sensitive to the first guess as the number of mass points m increases. This is due to the fact that optimization of the capacity given that the number of mass points m is fixed, is not a concave function. This accordingly, makes the problem computationally demanding with m.

V. CONCLUSIONS

In this paper, we studied the capacity of a complex AWGN channel in the presence of a nonlinear energy harvester at the receiver. Motivated by the two nonlinear models of the energy harvester introduced in the literature, two sets of constraints were studied. First, we considered the capacity problem under AP, PA and RDP constraints. We showed that the capacity
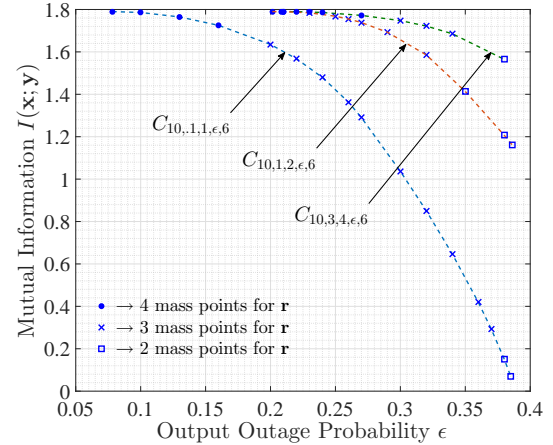
under the AP and RDP constraint, is the same as the capacity under AP constraint that can be either achieved or approached arbitrarily. In line with the similar results in the literature, we showed that including the PA constraint causes the amplitude of the optimal input to be discrete with a finite number of mass points. Next, we considered the capacity problem under AP, PA and OOP constraints, where similar results were derived.

REFERENCES

[1] L. R. Varshney, "Transporting information and energy simultaneously," in *IEEE Int. Sym. Inf. Theory*, Jul. 2008, pp. 1612–1616.
[2] M. S. Trotter, J. D. Griffin, and G. D. Durgin, "Power-optimized waveforms for improving the range and reliability of rfid systems," in *IEEE Int. Conf. RFID*, Apr. 2009, pp. 80–87.
[3] B. Clerckx and E. Bayguzina, "Waveform design for wireless power transfer," *IEEE Trans. Sig. Processing*, vol. 64, no. 23, Dec. 2016.
[4] B. Clerckx, "Wireless information and power transfer: Nonlinearity, waveform design and rate-energy tradeoff," *IEEE Trans. Sig. Processing*, vol. 66, no. 4, Feb. 2018.
[5] M. Varasteh, B. Rassouli, and B. Clerckx, "Wireless information and power transfer over an awgn channel: Nonlinearity and asymmetric gaussian signaling," in *IEEE Inf. Theory Workshop (ITW)*, Nov. 2017.
[6] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for swipt systems," *IEEE Commu. Letters*, vol. 19, no. 12, Dec. 2015.
[7] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, Jul. and Oct. 1948.
[8] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Inform. Control*, vol. 18, 1971.
[9] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature gaussian channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 1060–1071, Jul. 1995.
[10] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.
[11] J. J. Fahs and I. C. Abou-Faycal, "Using hermite bases in studying capacity-achieving distributions over awgn channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5302–5322, Aug. 2012.
[12] M. Varasteh, B. Rassouli, and B. Clerckx, "On capacity-achieving distributions over complex awgn channels under nonlinear power constraints and their applications to swipt," *CoRR*, vol. abs/1712.01226, 2017. [Online]. Available: http://arxiv.org/abs/1712.01226
[13] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: John Wiley & Sons, Inc., 1968.
[14] D. Luenberger, *Optimization by Vector Space Methods*. New York: John Wiley & Sons, Inc, 1969.