

# Information Energy Capacity Region for SWIPT Systems over Rayleigh-Fading Channels

Nizar Khalfet, *Member, IEEE*, and Ioannis Krikidis, *Fellow, IEEE*

**Abstract**—In this paper, we study the fundamental limits of simultaneous information and power transfer over a Rayleigh-fading channel, where the channel input is constrained to peak-power (PP) constraints that vary in each channel use by taking into account high-power amplifier (HPA) nonlinearities. In particular, a three-party communication system is considered, where a transmitter aims simultaneously conveying information to an information receiver and delivering energy to an energy harvesting receiver. For the special case of static PP constraints, we study the information-energy capacity region and the associated input distribution under: a) average-power and PP constraints at the transmitter, b) an HPA nonlinearity at the transmitter, and c) nonlinearity of the energy harvesting circuit at the energy receiver. By extending Smith's mathematical framework [1], we show that the optimal input distribution under those constraints is discrete with a finite number of mass points. We show that HPA significantly reduces the information energy capacity region. In addition, we derive a closed-form expression of the capacity-achieving distribution for the low PP regime, where there is no trade-off between information and energy transfer. For the case with time-varying PP constraints, we characterize the optimal input distribution of this channel by using Shannon's coding scheme. Specifically, we numerically study a particular scenario for the time-varying PP constraints, where the PP constraint probabilistically is either zero or equal to a non-zero constant.

**Index Terms**—SWIPT, wireless power transfer, high-power amplifier, optimal input distribution, information-energy capacity region, peak-power constraint.

## I. INTRODUCTION

Simultaneous wireless information and power transfer (SWIPT) is a technology that exploits the duality of the radio frequency (RF) signals, which can carry both information and energy [2] through appropriate co-design and engineering. The idea of wireless power transfer (WPT) was first proposed by Tesla in the 20-th century [3], and now presents a promising solution for modern communication systems such as low-power short-range communication systems, sensor networks, machine-type networks, and body-area networks [4]. The notion of the information-energy capacity region for SWIPT systems, was first formalized by Varshney [5] in the context of point-to-point scenarios. This work has been extended in [6] for a parallel links point-to-point channel. More recent works study the integration of SWIPT to more complex network topologies e.g., multiple access channel [7], interference channel [8], multiple-input multiple-output [9], multiple-antenna cellular networks [10], etc. A comprehensive

overview of existing results in SWIPT for various fundamental multi-user channels is presented in [11].

The design of the WPT component is crucial in order to characterize SWIPT systems. Most of the literature assumes simple linear models for the RF energy harvester (EH) receiver [9], [12] to simplify analysis. However, one of the main particularities of a SWIPT network is that the WPT channel is highly nonlinear (in contrast to the linear information transfer channel). Recent studies take into account the nonlinearity of the rectification circuit, and study the impact of waveform design and/or input distribution on the achieved information energy capacity region. For instance, the work in [13] models the rectifier behavior and introduces a mathematical framework to design waveforms that exploit nonlinearity. This observation introduces a relevant question for SWIPT networks: "what is the fundamental limit of a SWIPT system over a Rayleigh-fading channel with a non-linear EH receiver and constant/time-varying PP constraints?" The problem was first formalized in [14] by considering a truncated Taylor expansion series approximation for the diode's characterization function over an additive white Gaussian noise (AWGN) channel. The authors have shown that the optimal input distribution is zero-mean complex Gaussian distribution, with an asymmetric power allocation for the real and the imaginary parts. However, a more general model was proposed in [15], by using the exact form of the diode's characteristic function. The authors have extended Smith's mathematical framework [1] and have shown that the optimal input distribution under the first the and second moments statistics as well as a peak power (PP) constraint at the transmitter, it is unique, discrete with a finite number of mass points. In [?], the authors have shown that the capacity-achieving distribution is discrete with a finite number of mass points for the class of Gaussian mixture noise channel with minimum and peak-power (PP) constraints.

On the other hand, experimental studies demonstrate that signals with high peak-to-average-power-ratio (PAPR) e.g., multi-sine, chaotic signals, white noise, etc [16], provide a higher direct-current (DC) output, in comparison to constant-envelope sinusoidal signals [17]. However, signals with high PAPR are more sensitive to high-power amplifier (HPA) nonlinearities, which significantly degrade the quality of the communication [18], [19]. With the exception of a few studies (e.g., [20]), existing works do not consider the effects of HPA on SWIPT performance and assume that the HPA operates always in the linear regime. In addition, most of the aforementioned studies on SWIPT systems focus on simple AWGN channels and therefore the impact of the channel fading has not been investigated. To the best of the

N. Khalfet and I. Krikidis are with Department of Electrical and Computer Engineering, University of Cyprus, Cyprus (e-mail: {khalfet.nizar, krikidis}@ucy.ac.cy).

This work received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 81981)

authors' knowledge, this is the first work that takes into account the effect of HPA's non-linearity on SWIPT systems over fading channels from an information theoretic standpoint.

In practice, energy is harvested from nature through different sources, such that solar, vibrations, wind, etc. Therefore, the energy available at the transmitters is not always a deterministic process and can be modeled as a random process that varies in time [21]. In such systems, the transmitter casually observes the arrived energy and then makes a decision on the code symbol transmitted. This setup is modeled as a state-dependent channel with causal state information available at the transmitter. The fundamental limits of the state-dependent channel was first introduced by Shannon in [22], where a capacity achieving coding scheme was proposed. By extending the framework of [1] and [22], the authors characterized the optimal input distribution over an AWGN channel with time-varying PP constraints by extending the alphabet of the codewords in accordance with the PP constraints. The later work was extended for a Gaussian multiple access channel in [23], where the authors have shown that the boundary of the capacity region is achieved by a discrete input distribution with a finite support. However, none of these works addresses the fundamental limits of the fading channel with time-varying PP constraints in a SWIPT context.

This paper studies the fundamental limits of SWIPT over a Rayleigh-fading channel, where the channel input is constrained to a PP varying constraint at each channel use; in addition, we take into account a memoryless HPA model at the transmitter and a non linear power transfer channel. The varying PP constraint induces a state-dependent channel with a perfect causal state information at the transmitter, since the amplitude process is observed by the transmitter. In particular, we consider a basic three-node SWIPT system, where a transmitter simultaneously sends data to an information receiver and power to an EH receiver through a Rayleigh-fading channel; we consider both average power (AP) and time varying PP constraints at the transmitter as well as a non-linear power transfer channel. For the special case of static PP constraints, we characterize the information energy capacity region and we show that the associated capacity-achieving input distribution is unique, discrete, with a finite number of mass points. For the general case with time varying PP constraints, we prove that the capacity achieving distribution has a finite support by using Shannon's coding scheme [22]. Our study generalizes the result for the capacity-achieving input distribution of a conventional discrete-memoryless Rayleigh-fading channel under AP, which has been studied in [24]. We show that HPA significantly reduces the information energy capacity region, while increasing the PP constraint enlarges the associated region. Finally, we study the optimal input distribution for the low PP regime, where a no trade-off between information and energy transfer is observed.

The technical contribution of this paper is twofold:

- The case with a single static PP constraint ( $M = 1$  e.g., the transmitter is connected to the power grid), mainly

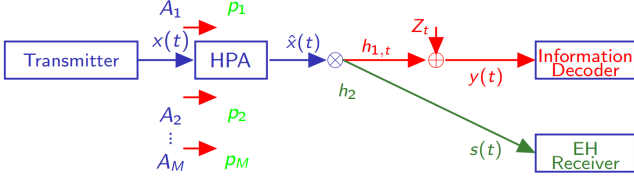
Notations	Description	Notations	Description
$n$	Number of channel uses	$X$	Random variable of the input signal
$M$	Cardinality of the states	$\mathcal{T}$	Random variable of the extended alphabet
$t$	Time index	$X^*$	Optimal input random variable
$P$	Average power constraint	$F^{(n)}$	Sequence of the the input distribution
$E_{\text{req}}$	Energy required at the EH	$F(\cdot)$	Probability distribution function of $X$
$h_{1,t}$	Channel fading for the information link	$I(\cdot)$	Mutual information as a function of the distribution $F$
$h_2$	Channel fading for the EH link	$F^*$	Optimal input probability distribution
$F^{(n)}$	Sequence of a distribution $F$	$i(x; F)$	mutual information density of $F$ evaluated at $x$
$F_N^*$	Optimal distribution with $N$ mass points	$\mathcal{CN}(0, \sigma_1^2)$	Circular complex Gaussian random variable with zero mean and $\sigma_1^2$ variance
$\text{Supp}(F)$	Support of a distribution $F$		

TABLE I: Summary of notation.

refers to scenarios where the transmitter is enforced to operate in the linear regime of the HPA; it is a way to control the HPA nonlinearities effects [25]. We show that the optimal input distribution that maximizes the information energy capacity is unique, discrete, with a finite number of mass points. We numerically study the impact of the HPA non-linearity on the information energy capacity region and we show that HPA significantly reduces the corresponding region. In addition, we propose a mathematical framework to study the optimal input distribution for the low PP regime, where there is not a trade-off between information and energy transfer.

- For scenarios with time-varying PP constraints at the transmitter (e.g., the energy arrives at the transmitter in a sporadic way according to a random process), we characterize the information energy capacity region by extending the results in [5]. Firstly, we construct a channel which has an input extended by the cardinality of the state alphabet and we optimize the input distribution by applying Shannon's strategy [22]. Then, we provide a necessary and sufficient condition for the optimal input distribution. For the sake of exposition, we deal with a scenario of a binary on-off energy arrival process, where either a constant amount of energy or no energy arrives.

*Notation:* In this paper, sets are denoted with uppercase calligraphic letters. Random variables are denoted by uppercase letters, e.g.,  $X$ . The realization and the set of the events from which the random variable  $X$  takes values are denoted by  $x$  and  $\mathcal{X}$ , respectively. The argument  $\mathbb{E}[X]$  denotes the expectation with respect to the distribution of a random variable  $X$ . The notation  $F$  denotes the probability distribution function of a random variable  $X$ , and  $F^*$  represents the optimal input distribution. The notation  $\text{Supp}(F)$  is the support of a distribution  $F$ , i.e.,  $\text{Supp}(F) = \{x \in \mathcal{X} \mid \int_{x \in \mathcal{O}} dF(x) > 0 \text{ for every open neighborhood } \mathcal{O} \text{ of } x\}$ .



**Fig. 1:** A SWIPT system over a fading channel with time varying PP constraints, HPA at the transmitter, and a non linear EH channel.

Table. I summarizes the key notation of the paper. The reminder of this paper is structured as follows. The system model is presented in Section II. In Section III, we characterize the information energy capacity region for the case of static PP constraints. For the time-varying PP constraint case, we study the optimal input distribution in Section IV. Finally, numerical results are presented in Section V.

## II. SYSTEM MODEL

Consider a three part communication system, where a transmitter aims simultaneously convey information to an information receiver (IR) and power to an EH receiver. The IR converts the received signal to the baseband to decode the transmit information, while the EH receiver harvests energy from the received RF signal. In each channel use, the transmitter inputs a pulse-amplitude modulated signal  $x(t) = \sum_{k=-\infty}^{\infty} x[k]p(t - kT)$ , with an average power  $P$ , where  $p(t)$  is the rectangular pulse shaping filter<sup>1</sup> (i.e.,  $p(t) = 1$  for  $0 < t \leq T$ ),  $T$  is the symbol interval, and  $x[k]$  is the information symbol at time index  $k$ , modeled as the realization of an independent and identically distributed (i.i.d) real random variable  $X$  with probability distribution  $F$ . We assume a normalized symbol interval  $T = 1$  and thus the measures of energy and power become identical and therefore are used equivalently. The system model is depicted in Fig. 1. The transmitted amplitude-modulated signal  $x(t)$  is subjected to nonlinearities induced by the HPA; the output of the nonlinear HPA can be written as  $\hat{x}[k] = d(x[k])$  (i.e., random variable  $\hat{X} = d(X)$ ), where  $d(\cdot)$  denotes the AM-to-AM conversion<sup>2</sup> which is given by the considered solid state power amplifier (SSPA) HPA model [18] i.e.,

$$d(r) = \frac{r}{\left[1 + \left(\frac{r}{A_s}\right)^{2\beta}\right]^{\frac{1}{2\beta}}}, \quad (1)$$

where  $A_s$  is the output saturation voltage, and  $\beta$  represents the smoothness of the transition from the linear regime to the saturation.

<sup>1</sup>This assumption is done for analytical tractability of the average energy harvested expression in (8), without loss of generality.

<sup>2</sup>Amplitude-to-amplitude modulation (AM-to-AM) represents one of the characteristics functions of the power amplifier, which denotes the amplitude distortion of the signal due to its non linearity effects. It is worth noting that AM-to-AM conversion can lead to several performance degradation. For the solid-state power amplifier (SSPA), the AM-to-AM conversion is given by equation (1).

### A. Information transfer

For the information transmission, we consider a memoryless discrete-time Rayleigh-fading channel [24]; the channel output at the receiver during the channel use  $t$  is given by

$$y(t) = h_{1,t}\hat{x}(t) + Z_t, \quad (2)$$

where  $\hat{x}(t)$  is the channel input induced by the HPA,  $y(t)$  is the channel output, and  $h_{1,t}$  and  $Z_t$  are independent complex circular Gaussian random variables distributed as

$$h_{1,t} \sim \mathcal{CN}(0, \sigma_1^2), \quad (3)$$

$$Z_t \sim \mathcal{CN}(0, \sigma_2^2). \quad (4)$$

By conditioning on the channel input,  $|Y|$  has a central chi-square distribution with two degree of freedom [24], as follows

$$P_{|Y|^2|\hat{X}}(\alpha|\hat{x}) = \frac{1}{\pi(\sigma_1^2|\hat{x}|^2 + \sigma_2^2)} \exp\left[\frac{-\alpha}{\sigma_1^2|\hat{x}|^2 + \sigma_2^2}\right]. \quad (5)$$

In order to obtain a single-letterization of the equivalent channel model given in (5), we define  $\tilde{Y} = \frac{|Y|^2}{\sigma_2^2}$  and  $\tilde{X} = \frac{|\hat{X}|^2}{\sigma_2^2}$ ; with this notation, an equivalent channel model is obtained with a nonnegative input  $\tilde{X}$ , a non negative output  $\tilde{Y}$  and a transition probability i.e.,

$$p(y|\hat{x}) = \frac{1}{1 + \hat{x}^2} \exp\left(-\frac{y}{1 + \hat{x}^2}\right). \quad (6)$$

### B. Power transfer

At the EH receiver, the contribution of the noise is assumed to be negligible<sup>3</sup>, hence, the representation of the received signal at the EH receiver (in the baseband equivalent) is given by

$$s(t) = h_2\hat{x}(t), \quad (7)$$

where  $h_2 \in \mathbb{R}$  is the channel fading for the link between the transmitter and the EH receiver; As for the wireless channel, we consider a conventional Rayleigh block fading channel for the information transfer [24]. We assume a quasi-static fading channel for the energy transfer, where the wireless channel between the transmitter and the EH receiver is assumed constant over a sufficiently long period of time [?]. The justification of this assumption is that in most of the cases, EH receivers are static and located close to the transmitter (i.e., in a strong line-of-sight link) and thus small-scale fading can be considered practically constant [?], [25]. We assume that the static channel  $h_2$  is known to both transmitter and receiver (a low rate feedback channel ensures this assumption [?]) in this work. For the average energy harvested, we adopt the nonlinear EH circuit model proposed in [15]. In particular, the EH circuit includes a rectenna, which is composed by an antenna and rectifier. The role of the rectifier is to convert the RF signal to a DC signal across to its load resistance. The rectifier is characterized by a single Shockley diode followed by a capacitor-based low pass filter. In order to ensure maximum power transfer, an ideal matching network is embedded to match

<sup>3</sup>In practical systems, the AWGN noise power is negligible compared to the signal power at the EH receiver [25]

the rectifier input independence to the antenna impedance. Although, several non-linear EH models have been proposed in the literature (e.g., a truncated Taylor series expansion of the diode's nonlinear characteristic equation is considered in [13]), we adopt a model that captures the exact form of the diode's characteristic equation [25]. Let  $\mathcal{E} : \mathbb{R} \rightarrow \mathbb{R}_+$  be the function that determines the average energy harvested [15], which is given by

$$\mathcal{E}(\hat{X}) = \mathbb{E} \left[ I_0 \left( \sqrt{2} B h_2 |\hat{X}| \right) \right], \quad (8)$$

where  $I_0(\cdot)$  is the modified Bessel function of the first kind and order zero and  $B$  is a constant that depends on the characteristics of the rectification circuit. The EH constraint is reduced as

$$\mathcal{E}(\hat{X}) \geq E_{\text{req}}. \quad (9)$$

The physical meaning of the energy constraint in (9) is that it denotes the minimum amount of energy which is required to operate/power the circuits of the EH-based device. In practical SWIPT applications, we deal with low-power devices (e.g., sensor nodes, RFID tags) where this constraint is low and thus can be supported by the received RF signals [?], [?]. On the other hand, from an information theory standpoint [5], by varying the energy constraint  $E_{\text{req}}$ , we are able to characterise the information-energy capacity region which is the main focus of our paper.

### C. Problem formulation

We consider a setup where the energy arrives from different natural sources [?], [?] as an independent and identically distributed (i.i.d.) ergodic random process [21], [26], [?]. Specifically, we assume that the energy is received in the system from a power source that delivers  $E_i$  units of energy at the  $i$  channel use, where  $E_1, \dots, E_n$  is the i.i.d energy arrival sequence in  $n$  channel uses. From an information-theoretic perspective, we adopt an arbitrary distribution for the energy arrival random variable  $E_i$ , such as Bernoulli process [?], Poisson arrival process [?] etc, without loss of generality. Regarding the feasibility of the proposed setup, we consider that the received energy is stored in a temporary storage device (e.g., supercapacitor [?]) and the harvest-then-use protocol is employed [?]. Specifically, the symbol transmission is performed in two phases: i) the storage phase where the storage device is charged with energy, and ii) the communication phase where the storage device is discharged and the energy is used for transmitting the code symbols. Therefore, the energy arrival process has the same discrete time index as the channel use. Hence we can ensure that the energy arrival can be observed causally by the transmitter as shown in Fig. 1. Let  $A$  be a PP constraint random variable with alphabet  $\mathcal{A} = \{a_1, a_2, \dots, a_M\}$ , where  $M$  is the cardinality of the alphabet of the energy arrivals. Let also  $\{A_k\}_{k=1}^\infty \in \mathcal{A}$  be the i.i.d. PP constraint process with

$$\Pr[A_k = a_i] = p_i, \quad \forall k, \quad (10)$$

with  $i \in \{1, 2, \dots, M\}$ . The realization of the PP constraint  $\{A_1, A_2, \dots, A_M\}$  is observed by the transmitter, and the code symbols must satisfy

$$X_k \leq A_k, \quad k \in \{1, 2, \dots, n\}, \quad (11)$$

where  $n$  is the number of channel uses. The consideration of discrete PP constraints is a well-known topic in studies dealing with the sporadic nature of the energy harvesting process, e.g., [26], [?]; by taking  $M$  (the cardinality of the amplitude constraint set) sufficient large, we can approximate the continuous case. Note that the variation of the amplitude at the transmitter is not known at the receiver; in this case the channel is state dependent with a causal state information available only at the transmitter. By applying the capacity scheme proposed by Shannon [22], We denote by  $T = [T_1, T_2, \dots, T_M]$  the extended input alphabet, that takes values into the set  $\mathcal{X} = [0, A_1] \times [0, A_2] \dots [0, A_M]$ , then the space of joint probability distribution function over  $\mathcal{X}$  is given by

$$\mathcal{F}_{A_1, A_2, \dots, A_M} \triangleq \left\{ F : \int_0^{A_1} \int_0^{A_2} \dots \int_0^{A_M} dF(t_1, t_2, \dots, t_M) = 1 \right\}. \quad (12)$$

The expression of the equivalent channel is given by

$$\begin{aligned} f(y|\hat{t}_1, \hat{t}_2, \dots, \hat{t}_M) &= \sum_{i=1}^M p_i p(y|\hat{t}_i) \\ &= \sum_{i=1}^M p_i \frac{1}{1 + \hat{t}_i^2} \exp\left(-\frac{y}{1 + \hat{t}_i^2}\right). \end{aligned} \quad (13)$$

In particular, the mutual information density for the equivalent channel (13) is given by

$$\begin{aligned} i(t_1, t_2, \dots, t_M; F) &= \int f(y|t_1, t_2, \dots, t_M) \\ &\times \log\left(\frac{f(y|t_1, t_2, \dots, t_M)}{f(y; F)}\right) dy. \end{aligned} \quad (14)$$

The optimal coding scheme for the state-dependent channel was investigated by Shannon [22]. This work has shown that the capacity of the state-dependent channel is equal to the capacity of an equivalent channel with an input alphabet extended by the cardinality of the state alphabet. Shannon strategy consists of generating a two dimensional codewords, where the number of columns is the block length and the number of rows is the cardinality of the state alphabet.

By using Shannon strategy, the information energy capacity region is obtained by maximizing the mutual information between  $T$  and  $Y$ . The optimal input distribution that achieves the information energy capacity region for the time-varying PP constraint fading channel is obtained by solving the following optimization problem

$$\begin{aligned} &\sup_{F \in \mathcal{F}_{A_1, A_2, \dots, A_M}} I(F) \\ &\text{subject to} \quad \sum_{i=1}^M p_i \mathbb{E}[T_i^2] \leq P, \\ &\quad \quad \quad \sum_{i=1}^M p_i \mathcal{E}(\hat{T}_i) \geq E_{\text{req}}, \end{aligned} \quad (15)$$

where

$$I(F) = \int_0^{A_1} \int_0^{A_2} \dots \int_0^{A_M} i(t_1, t_2, \dots, t_M; F) dF(t_1, t_2, \dots, t_M). \quad (16)$$

#### D. Smith's framework [1]

The capacity of PP-constrained Gaussian channel was first introduced by Smith in [1], where it was shown that the optimal input distribution that achieves the information capacity is discrete, unique, with a finite number of mass points. The basic steps of the framework are summarized as follows:

- 1) *Step 1*: Prove that the capacity achieving distribution is unique by showing that the space of distribution functions is concave in a weak topology.
- 2) *Step 2*: The dual problem is given by using the Lagrangian theorem [1], since the basic optimization is valuable for non constrained functions.
- 3) *Step 3*: Provide a necessary and sufficient condition for the capacity achieving distribution by using the optimization theorem.
- 4) *Step 4*: Prove that the optimal input distribution is discrete through contradiction, by using the identity theorem for analytic complex functions.

### III. SWIPT FOR STATIC PP CONSTRAINT

In this section, we focus on the special case where the PP constraint at the transmitter is static/constant, i.e.,  $\mathcal{A} = \{a_1\}$ . This case refers to SWIPT scenarios, where the transmitter has a constant power supply e.g., it is connected to the power grid. Hence the formulation in (15) is reduced to an one-dimensional optimization problem, and the input distribution is subjected to

$$|T_k| = |T| = |X| < A_1 = A, \quad \forall k. \quad (17)$$

The conditional probability in (13) is reduced to

$$p(y|\hat{x}) = \frac{1}{1 + \hat{x}^2} \exp\left(-\frac{y}{1 + \hat{x}^2}\right). \quad (18)$$

The main objective is to maximize the average mutual information between  $X$  and  $Y$  subject to both AP and PP constraints on the transmit symbols  $X$ , and a minimum harvested power constraint at the EH receiver. The optimization problem in (15) is reduced as

$$\begin{aligned} \sup_{F \in \mathcal{F}_A} \quad & I(F) = \int \int p(y|x) \log \frac{p(y|\hat{x})}{p(y; F)} dy dF(x), \\ \text{subject to} \quad & \mathbb{E}[X^2] \leq P, \\ & \mathcal{E}(\hat{X}) \geq E_{\text{req}}, \end{aligned} \quad (19)$$

where  $\mathcal{F}_A$  is the set of all input distributions that satisfies the PP constraint, i.e.,

$$\mathcal{F}_A = \left\{ F \in \mathcal{F}; \int_0^A dF(x) = 1 \right\}. \quad (20)$$

The mutual information between  $X$  and  $Y$ , as well the AP and PP constraints and the minimum harvested energy required at the EH, could be expressed in function of the distribution

$F$ . Denote by  $g_1 : F \rightarrow \mathbb{R}$  and  $g_2 : F \rightarrow \mathbb{R}$  the following functions

$$I(F) \triangleq \int_0^A i(x; F) dF(x), \quad (21)$$

$$g_1(F) \triangleq \int_0^A x^2 dF(x) - P, \quad (22)$$

$$g_2(F) \triangleq E_{\text{req}} - \int_0^A \mathcal{E}(\hat{x}) dF(x). \quad (23)$$

Denote by  $\Omega$  the set of the constrained input distribution, such that

$$\Omega = \left\{ F \in \mathcal{F}; \int_0^A dF(x) = 1; g_i(F) \leq 0; i \in \{1, 2\} \right\}, \quad (24)$$

then the optimization problem in (19) can be written simply as  $C = \sup_{F \in \Omega} I(F)$ .

#### A. Discreteness of the optimal input distribution

In this section, we study the properties of the capacity achieving distribution, i.e., the solution of the optimization problem in (19). By extending the mathematical framework in [1], Theorem 1 establishes the existence and the uniqueness of the optimal input distribution (first step of Smith's framework). By using the Lagrangian Theorem, the dual equivalent problem is given by Corollary 1 (second step of Smith's framework). Furthermore, we give necessary and sufficient conditions for the optimal input distribution in Corollary 2 (third step of Smith's framework). Finally, we show that the capacity-achieving input distribution is discrete in Theorem 3.

**Theorem 1.** *The capacity  $C$  is achieved by a unique input distribution  $F^*$ , i.e.,*

$$C = \sup_{F \in \Omega} I(F) = I(F^*). \quad (25)$$

*Proof:* The proof is presented in Appendix A. ■

**Remark 1.** *Note that for the case where the EH constant is omitted, the problem is reduced to the conventional problem of information capacity over a Rayleigh fading channel [24], where the optimal input distribution is discrete with a finite number of mass points (even without considering the AP constraint).*

To establish a necessary and sufficient condition for the optimal input distribution, we refer to the optimization theorem [1], which determines the optimal elements for a non-constrained convex space. In the following Corollary, we transform the optimization problem in (25), to the dual problem so we are able to apply the basic optimization theorem presented in [1].

**Corollary 1.** *The strong duality holds for the optimization problem in (25), i.e., there exist constants  $\lambda_1 \geq 0$  and  $\lambda_2 \geq 0$  such that*

$$C = \sup_{F \in \mathcal{F}_A} I(F) - \lambda_1 g_1(F) - \lambda_2 g_2(F), \quad (26)$$

*Proof:* The proof is presented in Appendix B. ■

The following Theorem establishes a necessary and sufficient condition on the optimal input distribution.

**Theorem 2.**  $F^*$  is the capacity achieving input distribution, if and only if,  $\forall F \in \Omega$ , there exist  $\lambda_1 \geq 0$  and  $\lambda_2 \geq 0$  such that

$$\int i(x; F^*) - C - \lambda_1 g_1(F) - \lambda_2 g_2(F) dF(x) \leq C - \lambda_1 P + \lambda_2 E_{\text{req}}. \quad (27)$$

*Proof:* Despite the similarity with proof in [24], the complete proof is fully described in Appendix C for the sake of completeness. ■

By using the condition in (27), we derive a necessary and sufficient condition for the optimal input distribution in the following corollary.

**Corollary 2.** Let  $\text{Supp}(F^*)$  be the points of support of a distribution  $F^*$ , then  $F^*$  is the optimal input distribution, if there exist  $\lambda_1 \geq 0$  and  $\lambda_2 \geq 0$ , such that

$$\lambda_1 (x^2 - P) - \lambda_2 (I_0(\sqrt{2}Bh_2\hat{x}) - E_{\text{req}}) + C - \int p(y|\hat{x}) \log \frac{p(y|x)}{p(y; F^*)} dy \geq 0, \quad (28)$$

for all  $x$ , with equality if  $x \in \text{Supp}(F^*)$ .

*Proof:* The proof is presented in Appendix D. ■

By using the invertible change of variables, i.e.,  $S = \frac{1}{1+x^2}$ , we have

$$p(y|s) = s \exp(-ys), \quad s \in \left[ \frac{1}{1+A^2}, 1 \right]. \quad (29)$$

It is worth noting, that the channel input is nonnegative as well as amplitude constrained, i.e.,  $x \in [0, A]$  and by using the fact that  $x \rightarrow \frac{1}{1+x^2}$  is a non-increasing function, we have

$$\frac{1}{1+A^2} \leq s \leq 1. \quad (30)$$

Now the following proposition derives a necessary and sufficient condition for the optimal random variable  $S^*$ .

**Proposition 1.**  $S^*$  is the optimal random variable in  $\left[ \frac{1}{1+A^2}, 1 \right]$ , if there exist  $\lambda_1 \geq 0$  and  $\lambda_2 \geq 0$ , such that

$$\lambda_1 \left( \frac{1}{s} - 1 - P \right) - \lambda_2 \left( I_0(\sqrt{2}Bh_2 \left( \sqrt{\frac{1}{s}} - 1 \right) - E_{\text{req}} \right) + C - \log s + 1 + \int_0^\infty s e^{-sy} \log p(y; F^*) dy \geq 0, \quad (31)$$

for all  $s \in \left[ \frac{1}{1+A^2}, 1 \right]$ , with equality if  $s \in \text{Supp}(S^*)$ .

In the following, we show that the equality in Proposition 1 can not be satisfied in a set that has an accumulation point<sup>4</sup>, hence the support of  $S^*$  must be discrete. The discreteness property of the optimal input distribution is given by Theorem 3.

**Theorem 3.** Let  $F^*$  be the optimal input distribution that achieves the capacity in (19);  $F^*$  is discrete with a finite number of mass points.

<sup>4</sup>A point  $x$  is an accumulation point (limit point) of a set  $\mathcal{A}$ , if every open neighborhood of  $x$  contains at least one point from  $\mathcal{A} \subset \mathcal{X}$  distinct from  $x$ .

**Remark 2.** Note that the optimization problem in (19) could be generalized to  $K$  identical EH receivers; in this case, the optimization problem is written as

$$\begin{aligned} \sup_{F \in \mathcal{F}} \quad & I(F) = \int \int p(y|x) \log \frac{p(y|x)}{p(y; F)} dy dF(x), \\ \text{subject to} \quad & \mathbb{E}[X^2] \leq P, \\ & \mathcal{E}(X) \geq E_{k, \text{req}}, \quad \forall k \in \{1, 2, \dots, K\}. \end{aligned} \quad (32)$$

It has been shown in [25] that if  $A < A_s$ , then at most one EH constraint is active. Hence, the result of Theorem 3 is applied for this case as well.

*Proof:* The proof is presented in Appendix E. ■

**Remark 3.** By using the fact that  $F^*$  is discrete with a finite number of mass points, the problem in (19) is reduced to the determination of the maximum for a function of a finite dimensional vector, where its components are the mass points and their locations [1]. An arbitrary input distribution  $F$  can be written as

$$F(x) = \sum_{i=1}^N q_i u(x - x_i), \quad (33)$$

where  $x \rightarrow u(x)$  is the unit step function. In practice, it would be interesting to impose a constraint on the number of mass points, i.e.,  $N < N_0$ ; since in practice, the constellation size of the transmitted signal is limited. In section V, we plot the information energy capacity region for different amplitude-shift keying (ASK) modulations. The original optimization problem in (15) is written as

$$\begin{aligned} \sup_{(x_1, x_2, \dots, x_N, q_1, q_2, \dots, q_N)} \quad & I(x_1, x_2, \dots, x_N, q_1, q_2, \dots, q_N), \\ \text{subject to} \quad & \sum_{i=1}^N q_i = 1, \\ & \sum_{i=1}^N q_i x_i^2 \leq P, \\ & \sum_{i=1}^N q_i \left[ I_0(\sqrt{2}Bh_2 x_i) \right] \geq E_{\text{req}}, \\ & N < N_0. \end{aligned} \quad (34)$$

Although the optimization problem in (19) is convex over all input distribution functions, we can not claim that the problem in (34) remains convex since it is parametrized by the mass point probability and its locations cite [24]. However, we have shown through the proof of Theorem 1, that the problem in (19) is convex for any distribution  $F \in \mathcal{F}$ ; thus the problem remains convex for the distribution  $F(x) = \sum_{i=1}^N q_i u(x - X_i)$ , where  $u(x)$  is the unit step function [15]. The problem in (34) can be solved numerically by using a numerical solver such as CVX [27]. More specifically, we discretize the interval  $[0, A]$ , with a sufficiently small step size  $\Delta x$  [15] to obtain the mass points set  $(x_1, x_2, \dots, x_N)$ . By fixing the symbol set, the optimization problem in (34) is convex with respect to the mass points  $(q_1, q_2, \dots, q_N)$ . Therefore, by using a very small step

$\Delta x \rightarrow 0$ , we obtain the global optimum, which is close to the optimal distribution. Finally, we check the optimality of the obtained input distribution by verifying the necessary and sufficient conditions of Corollary 2.

### B. Properties of the mass points

We give some insights on the behavior of the optimal input distribution  $F^*$  with respect to both the PP and the AP constraints. It has been shown that the optimal input distribution is discrete, therefore the CDF is determined by the vectors  $\mathbf{q}, \mathbf{x}$ , and the number of the mass points  $N$ . Specifically, the location of the mass points is given by

$$\mathbf{x} = (x_1, \dots, x_N), \quad (35)$$

and the weights associated with the mass points,

$$\mathbf{q} = (q_1, \dots, q_N). \quad (36)$$

Without loss of generality, we assume that  $x_1 < x_2 \dots < x_n < A$ . In addition, we assume that  $F_N^*$  is the optimal input distribution and is characterized by the triplet  $(\mathbf{q}^*, \mathbf{x}^*, N)$ , which is the solution of the optimization problem in (19). The conditions which are satisfied by the optimal input distribution for some  $\lambda_1 \geq 0$  and  $\lambda_2 \geq 0$  are

$$\begin{aligned} i(x; F_N^*) &\leq C + \lambda_1 (x^2 - P) - \lambda_2 (\mathcal{E}(x) - E_{\text{req}}), \text{ for } x \in [0, A], \\ i(x_i^*; F_N^*) &= C + \lambda_1 (x_i^{*2} - P) - \lambda_2 (\mathcal{E}(x_i^*) - E_{\text{req}}). \end{aligned} \quad (37)$$

Denote by  $g$  the following function i.e.,

$$g(w, F_N^*) = i(w; F_N^*) - \lambda_1 (w^2 - P) + \lambda_2 (\mathcal{E}(w) - E_{\text{req}}) \quad (38)$$

By using the fact that a point of support (except  $A$ ) is a local maximum for the function  $g(x, F_N^*)$ , it follows

$$\left. \frac{\partial g(w, F_N^*)}{\partial w} \right|_{w=x_i} = 0. \quad (39)$$

Unfortunately, a closed form solution for (39) seems unlikely for our setup. However, in the next remark, we are able to characterize a particular mass point for the optimal input distribution.

**Remark 4.** The input distribution has a necessary mass point at zero; by contradiction, we assume that  $0 < x_1$ , then by following similar analytical steps with [24], it can be shown that

$$\left. \frac{\partial i(w, F_N^*)}{\partial w} \right|_{w=x_1} < 0. \quad (40)$$

Furthermore, it can be shown that, for  $x_1 > 0$  we have

$$\left. - \frac{\partial \left( \lambda_1 (w^2 - P) - \lambda_2 (\mathcal{E}(w) - E_{\text{req}}) \right)}{\partial w} \right|_{w=x_1} < 0. \quad (41)$$

By using (40) and (41), then we have

$$\left. \frac{\partial g(w, F_N^*)}{\partial w} \right|_{w=x_1} < 0. \quad (42)$$

Hence,  $x_1$  is not a point of support for the function  $g(x, F_N^*)$ .

In the following, we study the behavior of the optimal input distribution at the transition point, where the binary distribution is no longer optimal by using the conjecture in [28]. This is a critical point in SWIPT systems, since the binary distribution maximizes both information and energy transfer simultaneously and therefore there is not a trade-off between them [20]. For simplicity, we assume that the AP constraint is active, for a low PP constraint, the optimal input distribution is binary and is characterized by  $\mathbf{x}^*$  and  $\mathbf{q}$ , i.e.,

$$\mathbf{x}^* = (0, x_1), \quad (43a)$$

$$\mathbf{q} = \left( 1 - \frac{P}{x_1^2}, \frac{P}{x_1^2} \right). \quad (43b)$$

Note that if

$$\frac{\partial I(Y, F_N^*)}{\partial x_1} > 0, \quad (44)$$

then we have  $x_1 = A$ . Let us assume that at the amplitude value  $\bar{A}$ , the binary distribution is not longer optimal. Thus at the amplitude value  $\bar{A} + \Delta \bar{A}$ , a new mass point appears denoted by  $x_2^* \in [0, A]$ , which satisfies

$$i(x_2^*; F_3^*) = C + \lambda_1 (x_2^{*2} - P) - \lambda_2 (\mathcal{E}(x_2^*) - E_{\text{req}}). \quad (45)$$

Let  $q$  the transition probability associated to the mass point  $x_2^*$ , then the optimal input distribution  $F_3^*$  is characterized by

$$\begin{aligned} q^* &= \left( 1 - \frac{P}{(x_1 + \Delta x_1)^2} - \left( 1 - \frac{x_2^*}{(x_1 + \Delta x_1)^2} \right) q, q, \right. \\ &\quad \left. \frac{P}{(x_1 + \Delta x_1)^2} - \frac{x_2^{*2} q}{(x_1 + \Delta x_1)^2} \right), \end{aligned} \quad (46)$$

$$\mathbf{x}^* = (0, x_2^*, x_1 + \Delta x_1). \quad (47)$$

**Remark 5.** Note that for low PP constraints, there is not a trade-off between information and energy transfer, since the optimal input distribution is binary and hence it maximizes both information and energy transfer simultaneously. By increasing the PP constraint, we show that there exist an amplitude value  $\bar{A}$ , in which the binary distribution is not longer optimal and hence, a trade-off between information and energy transfer is observed. The transition point  $\bar{A}$  at the low PP regime, satisfies the following condition

$$I(F_3^*(\bar{A} + \Delta \bar{A})) > I(F_2^*(\bar{A} + \Delta \bar{A})). \quad (48)$$

Hence by choosing  $\Delta$  small enough ( $\Delta \bar{A} \rightarrow 0$ ), we obtain a sufficient condition for the transition point  $\bar{A}$ , where the binary distribution is no longer optimal by using the following conjecture in [29] i.e., the number of mass points increases at most by one when the PP constraint is slightly increases. A more rigorous proof that avoids this conjecture is given in [?], and will be considered in our future work. It is worth noting that the sufficient condition in (48) offers new insights and allows the numerical computation of the transition point as shown in Fig. 2.

**Remark 6.** It was worth noting that the authors in [25] propose a sub-optimal distribution for any PP regime by examining the extreme cases of wireless information transfer (WIT) and wireless power transfer (WPT). In contrast with this work, the optimal input distribution for WIT for the Rayleigh-fading channel is an open problem in the literature and only the properties of the distribution are investigated [24]. Hence, we could not provide a sub-optimal distribution as discussed in [25] for our problem. However in section III-B, we have studied the binary optimal distribution that maximizes both information and energy for low peak power amplitude regimes.

#### IV. SWIPT FOR TIME-VARYING PP CONSTRAINTS

In this section, we elaborate the structure of the capacity achieving input distribution for the general system model shown in Fig. 1. For simplicity and without loss of generality, we assume that the PP constraint takes only two values  $a_1$  and  $a_2$  in a probabilistic way. Note that in the case where the amplitude variation is available at the transmitter, coding should be performed according the minimum amplitude. In this case, by applying Shannon's strategy [22], the extended input is reduced to  $\mathbf{T} = [X_1, X_2]$  that takes values into the set  $\mathcal{X} = [0, A_1] \times [0, A_2]$ , then the space of the admissible joint probability distribution function over  $\mathcal{X}$  is given by

$$\mathcal{F}_{A_1, A_2} \triangleq \left\{ F : \int_0^{A_1} \int_0^{A_2} dF(t_1, t_2) = 1 \right\}. \quad (49)$$

The expression for the equivalent channel is given by

$$\begin{aligned} f(y|\hat{t}_1, \hat{t}_2) &= p_1 p(y|\hat{t}_1) + p_2 p(y|\hat{t}_2) \\ &= p_1 \frac{1}{1 + \hat{t}_1^2} \exp\left(-\frac{y}{1 + \hat{t}_1^2}\right) + p_2 \frac{1}{1 + \hat{t}_2^2} \exp\left(-\frac{y}{1 + \hat{t}_2^2}\right). \end{aligned} \quad (50)$$

In this case, the mutual information that we aim to maximize is given by

$$I(\mathbf{T}; Y) = \int_0^{A_2} \int_0^{A_1} \int f(y|t_1, t_2) \log\left(\frac{f(y|t_1, t_2)}{f(y; F)}\right) dy dF, \quad (51)$$

where

$$f(y; F) = \int_0^{A_1} \int_0^{A_2} f(y|t_1, t_2) dF(t_1, t_2), \quad (52)$$

with

$$f(y|t_1, t_2) = p_1 p(y|t_1) + p_2 p(y|t_2), \quad (53)$$

and the equivalent of the mutual information density is given by

$$i(t_1, t_2; F) = \int f(y|t_1, t_2) \log\left(\frac{f(y|t_1, t_2)}{f(y; F)}\right) dy. \quad (54)$$

The original optimization problem in (15) is reduced to

$$\begin{aligned} &\sup_{F \in \mathcal{F}_{A_1, A_2}} I(F), \\ &\text{subject to } p_1 \mathbb{E}[T_1^2] + p_2 \mathbb{E}[T_2^2] \leq P, \\ &\quad p_1 \mathbb{E}[\mathcal{E}(\hat{T}_1)] + p_2 \mathbb{E}[\mathcal{E}(\hat{T}_2)] \geq E_{\text{req}}, \end{aligned} \quad (55)$$

where

$$I(F) = \int \int \int f(y|t_1, t_2) \log \frac{f(y|t_1, t_2)}{f(y; F)} dy dF(t_1, t_2). \quad (56)$$

In the following, we prove that the optimal input distribution that solves the optimization problem in (55) has a support set of a finite cardinality. First, by using similar arguments as the static PP constrained case, we can establish the uniqueness of the optimal input distribution; this follows from the following claims

- 1)  $I(\mathbf{t}; F)$  is a concave function and weakly differentiable in  $\mathcal{F}_{A_1, A_2}$ .
- 2)  $\mathcal{F}_{A_1, A_2}$  is a convex and compact space.
- 3) The PP and AP constraints are linear.
- 4)  $i(t_1, t_2; F)$  is continuous and analytical.

By following a similar procedure as the static PP constraint case, the following theorem characterizes a necessary and sufficient condition for the optimal input distribution.

**Theorem 4.**  $F^*$  is the optimal input distribution, if there exist  $\lambda_1 \geq 0$  and  $\lambda_2 \geq 0$ , such that

$$\begin{aligned} &\lambda_1 (p_1 t_1^2 + p_2 t_2^2 - P) - \lambda_2 (p_1 I_0(\sqrt{2} B h_2 \hat{t}_1) + p_2 I_0(\sqrt{2} B h_2 \hat{t}_2) \\ &- E_{\text{req}} + C - \int p(y|t_1, t_2) \log \frac{p(y|t_1, t_2)}{p(y; F^*)} dy) \geq 0, \end{aligned} \quad (57)$$

for all  $(t_1, t_2)$ , with equality if  $(t_1, t_2) \in \text{Supp}(F^*)$ .

By using the invertible change of variables,

$$S \triangleq \left( \frac{1}{1 + T_1^2}, \frac{1}{1 + T_2^2} \right), \quad (58)$$

a simpler necessary and sufficient condition on the optimal random variable  $S^* = \left( \frac{1}{1 + T_1^{*2}}, \frac{1}{1 + T_2^{*2}} \right) = (S_1, S_2)$  is derived in the following corollary.

**Corollary 3.**  $S^*$  is the optimal random variable, if and only if, there exist  $\lambda_1 \geq 0$  and  $\lambda_2 \geq 0$ , such that

$$\begin{aligned} &\lambda_1 \left( p_1 \left( \frac{1}{s_1} - 1 \right) + p_2 \left( \frac{1}{s_2} - 1 \right) - a \right) - \lambda_2 \left( p_1 \mathcal{E} \left( \sqrt{\frac{1}{s_1} - 1} \right) \right. \\ &\quad \left. + p_2 \mathcal{E} \left( \sqrt{\frac{1}{s_2} - 1} \right) - b \right) + C \\ &\quad + \int_0^\infty \sum_{i=1}^2 p_i s_i e^{-s_i y} \log \frac{\sum_{i=1}^2 p_i s_i e^{-s_i y}}{p(y)} dy \geq 0, \end{aligned}$$

for all  $(s_1, s_2)$ , with equality if  $(s_1, s_2) \in \text{Supp}(S^*)$ .

The next natural task is to prove that the optimal input distribution  $S^*$  has a finite support, by applying the identity theorem on the necessary and sufficient condition in Corollary 3. However, the identity theorem is not appropriate for the multi-dimensional scenarios. Hence, we could not show that  $S^*$  has a finite support as discussed in [26]. Specifically, the authors in [?] have proven that the application of the identity theorem from the complex analysis in several variables is not admissible by showing a simple example. From a practical standpoint, these potential difficulties do not arise as shown in the following remark.



**Remark 7.** Assuming that  $S^*$  has an infinite support, a numerical optimization of the problem in (55) yields a suboptimal input distribution and a lower bound on the information energy capacity region. By following the same procedure as [?], we increase the number of mass points, such that the lower bound can be further refined. Specifically, we fix the cardinality of the support as  $|\text{Supp}(F)| = 2$ , in this case, the optimal input distribution has two mass points characterized by  $(t_1, t_2) = (0, 0)$  and  $(t_1, t_2) = (a_1, a_2)$ . Then, we increase the number of probability mass points by one at the time to improve the mutual information until the improvement is less than  $10^{-3}$ .

#### A. On-off energy arrivals

We consider the special case of on-off energy arrival, where at each channel use, either a constant amount of energy or a zero energy arrives at the transmitter [26] in a probabilistic way. The alphabet of the amplitude random variable for this scenario is reduced to  $\mathcal{A} = \{0, a_2\}$  where  $p_2$  is the probability that the transmitter receives energy. The optimization problem in (59) is over the one dimensional distribution  $F(0, x_2)$  and is given by

$$\begin{aligned} \sup_{F \in \mathcal{F}_{A_2}} \quad & I(F) = \int \int f(y|t_2) \log \frac{f(y|t_2)}{f(y; F)} dy dF(t_2), \\ \text{subject to} \quad & p_2 \mathbb{E}[T_2^2] \leq P, \\ & p_2 \mathbb{E}[\mathcal{E}(\hat{T}_2)] \geq E_{\text{req}}, \end{aligned} \quad (59)$$

where

$$f(y|t_2) = (1 - p_2)p(y) + p_2p(y|t_2). \quad (60)$$

**Remark 8.** Similar to the static PP constraint case, for a low PP regime, i.e.,  $A \leq \bar{A}$ , the optimal input distribution  $F^*$  is symmetric and binary with two mass points at  $(0, A)$ . Hence, there is not a trade-off between the information and the energy transfer. For the special case with  $p_2 = 1$ , the problem reduces to the static PP-constraint Rayleigh fading channel.

### V. NUMERICAL RESULTS

We numerically evaluate the information-energy capacity region by using a numerical solver such CVX [27]. For the scenario with a static PP constraint, Fig. 2 shows the information-energy capacity region for different PP constraints. The corresponding region is obtained by solving the optimization problem in (19). A trade-off is observed between the information rate transmitted to the information receiver and the energy delivered to the EH receiver; this trade-off becomes evident since for higher EH constraints, the transmitter selects a symbol with a higher amplitude, which on the other hand, it degrades the information transfer performance. Another interesting remark is that for low PP constraints, there is not a trade-off between the two objectives. The optimal input distribution for this regime is binary, hence it maximizes both information and energy transfer simultaneously (Remark 5). In Fig. 3, we examine the special case of small PP constraints i.e.,  $A = 2V$ ; it can be seen the optimal input distribution

is binary with two mass points (43). This observation is also inline with Remark 5.

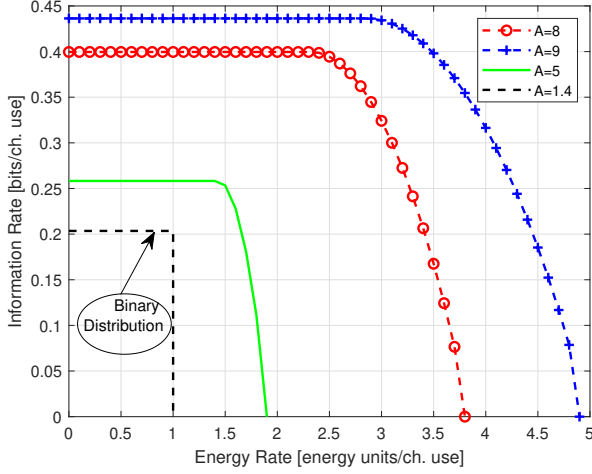
Fig. 4 highlights the effect of the HPA non-linearity on information-energy capacity region. There is a gap between the two regions; this is mainly due to the negative effect of the HPA. Fig. 5 plots the information energy capacity region for ASK modulations, which is obtained by solving the optimization problem in Remark 3 for different  $N_0$  values. It can be seen that by increasing the alphabet size, a smaller gap is observed between the rate achieved by the finite alphabet and the rate achieved by the optimal input distribution. For the sake of presentation, in our numerical study, we consider a binary scenario which is also inline with the work in [26]. The key reason is that for multiple peak power constraints, the numerical results incorporate multiple dimension curves, which make the presentation more complicated. Specifically, Fig. 6 deals with the time-varying PP constraint case under an on-off energy arrival, where the state information is available causally for different values of  $p_2$ . We observe that an increase of  $p_2$ , enlarges the information energy capacity region. Finally, Fig. 7 shows the impact of  $\beta$  (represents the smoothness of the transition from the linear regime to the saturation) on the information energy capacity region. It can be seen that by increasing  $\beta$ , i.e., the nonlinear transition regime (below saturation) of the HPA is linearized, and the information energy capacity region is enlarged.

### VI. CONCLUDING REMARKS

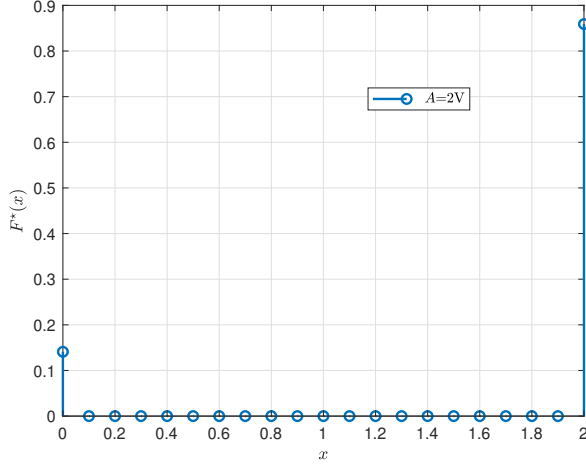
In this paper, we studied the fundamental limits of a SWIPT system over a Rayleigh- fading channel with time-varying constraints, non-linear EH, and by taking into account the non-linearity imposed by the HPA. For the special case of static PP constraints, we proved that the input distribution that maximizes the information-energy capacity region is unique, discrete, with a finite number of mass points. We have shown that the information energy capacity region increases by relaxing the PP constraint, while the HPA significantly degrades the performance of both objectives. Also, we proposed a mathematical framework to study the capacity achieving distribution for low PP constraints, where there is not a trade-off between information and energy transfer. For the case with time-varying constraints, we have derived a necessary and sufficient condition for the optimal input distribution by using Shannon's coding scheme. Finally, we have studied the optimal input distribution for the particular scenario of on-off energy arrival. In our future work, we intend to concentrate on the fundamental limits of SWIPT for more general settings e.g., Rician fading channel.

#### APPENDIX A PROOF OF THEOREM 1

Despite a great deal of similarity with the proof in [24], [?], [?], a sketch of the proof is presented for the sake of completeness. To prove the existence and the uniqueness of the optimal input distribution, it suffices to prove that the problem in (19) is convex, i.e., we need to prove the following points



**Fig. 2:** Information-energy capacity region for the static case with different PP constraints,  $\beta = 1$ ,  $P = 30$  dBW, and  $\sigma_1^2 = \sigma_2^2 = -80$  dBm.



**Fig. 3:** Optimal input distribution for the static case with a low PP constraint;  $\beta = 1$ ,  $P = 30$  dBW, and  $\sigma_1^2 = \sigma_2^2 = -80$  dBm.

- (i) The mutual information as function of the distribution  $F$ , i.e.  $I : F \rightarrow I(F)$  is a weak continuous function.
- (ii)  $I : F \rightarrow I(F)$  is strictly concave function.
- (iii) The set  $\Omega$  is compact.

*Continuity:* By using the definition of the convergence, we need to show that for any sequence of probability distributions  $F^{(n)}$  that converges to  $F$ ,  $I(F^{(n)})$  converges to  $I(F)$ . The mutual information for a specific distribution  $F$  could be written as

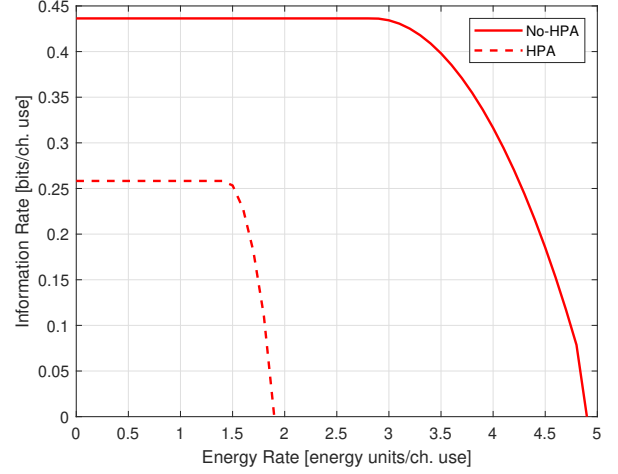
$$I(F) = h_Y(F) - h_{Y|X}(F), \quad (61)$$

where

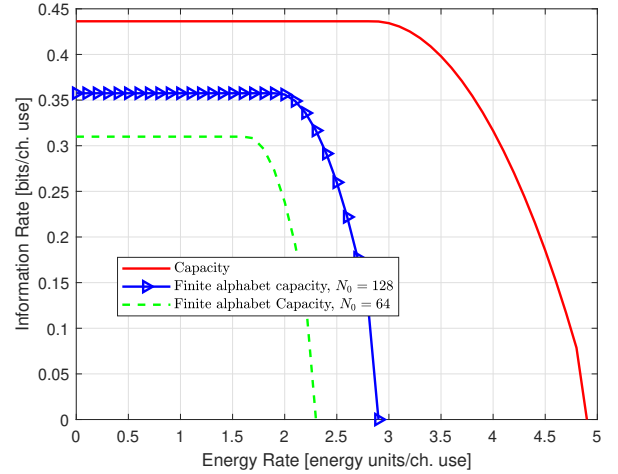
$$h_Y(F) \triangleq - \int_{-\infty}^{\infty} p(y; F) \log p(y; F) dy, \quad (62)$$

$$h_{Y|X}(F) \triangleq - \int_{-\infty}^{\infty} \int_{-A}^A p(y|x) \log p(y|x) dF(x) dy, \quad (63)$$

are the entropy of the random variables  $X$  and  $X|Y$  respectively. First, we need to show that the function  $F \rightarrow h_Y(F)$



**Fig. 4:** Effects of the HPA on the Information-energy capacity region;  $A = 5$ ,  $\beta = 1$ ,  $B = 0.5$ ,  $P = 30$  dBW, and  $\sigma_1^2 = \sigma_2^2 = -80$  dBm.



**Fig. 5:** Information energy capacity region for different finite alphabets  $N_0$ ;  $A = 5$ ,  $\beta = 1$ ,  $B = 0.5$ ,  $P = 30$  dBW, and  $\sigma_1^2 = \sigma_2^2 = -80$  dBm.

is weak continuous. Let  $F^{(n)} \rightarrow F$ , we need to show that  $h_Y(F^{(n)}) \rightarrow h_Y(F)$  by establishing the following equality

$$\lim_n h_Y(F^{(n)}) = - \lim_n \int p(y; F^{(n)}) \log p(y; F^{(n)}) dy \quad (64)$$

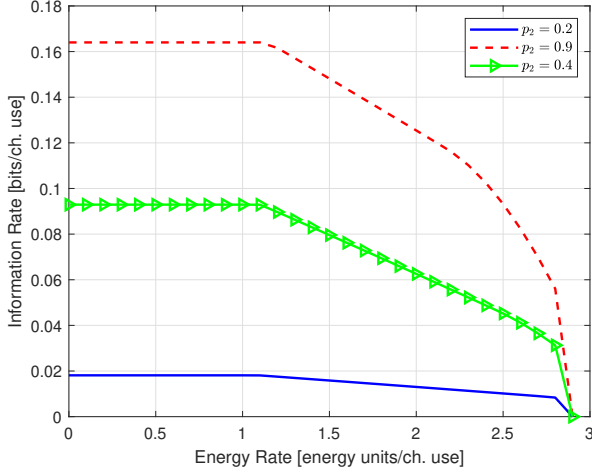
$$= - \int \lim_n p(y; F^{(n)}) \log p(y; F^{(n)}) dy \quad (65)$$

$$= - \int p(y; F) \log p(y; F) dy \quad (66)$$

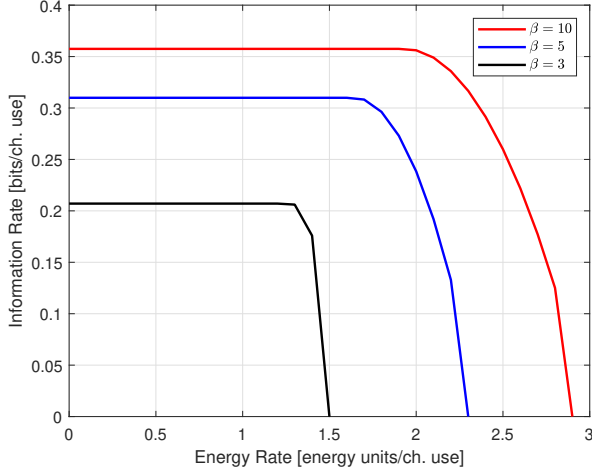
$$= h_Y(F), \quad (67)$$

where (64) and (67) are from definition; (65) follows from Lebesgue dominated convergence Theorem [30]; (66) follows from the continuity of the function  $x \rightarrow x \log x$ . The proof of continuity of  $F \rightarrow h_{Y|X}(F)$  follows the same arguments.

*Concavity, compactness:* Note that  $I(F) = h_Y(F) - h_{Y|X}(F)$  and since the function  $P_Y \rightarrow h_Y(P_Y)$  is a strictly concave function, and by using the fact that  $F \rightarrow P_Y(F)$  is a linear function, it holds that  $F \rightarrow h_Y(F)$  is a strictly concave. In addition, the function  $F \rightarrow h_{Y|X}(F)$  is a linear function, then



**Fig. 6:** Effect of  $p_2$  on the Information-energy capacity region for the time-varying case;  $A = 7$ ,  $\beta = 1$ ,  $B = 0.5$ ,  $P = 30$  dBW, and  $\sigma_1^2 = \sigma_2^2 = -80$  dBm.



**Fig. 7:** Information energy capacity region for different  $\beta$ ;  $A = 5$ ,  $B = 0.5$ ,  $P = 30$  dBW, and  $\sigma_1^2 = \sigma_2^2 = -80$  dBm.

$I : F \rightarrow I(F)$  is a strictly concave function. The proof of the compactness of  $\Omega$  is similar as [24].

#### APPENDIX B PROOF OF COROLLARY 1

Define a linear vector space  $X$ , a normed space  $Z$ , a convex subset of  $X$  denoted by  $\mathcal{F}$ , and a positive cone in  $Z$  that contains an interior point. Let  $f$  be a real-valued concave functional on  $\mathcal{F}$  and  $g$  is a convex mapping from  $\mathcal{F}$  to  $Z$ . Assume the existence of a point  $F_1 \in \mathcal{F}$ , for which  $g(F_1) < 0$  (Slater's condition). Let

$$C = \sup_{F \in \mathcal{F}, g(F) \leq 0} f(F), \quad (68)$$

and assume  $C$  is finite. Then by the Lagrangian theorem [1], there is an element  $z_0^* \in Z$  such that

$$C = \sup_{F \in \mathcal{F}} \{f(F) - \langle g(F), z_0^* \rangle\}, \quad (69)$$

where  $\langle \cdot, \cdot \rangle$  denotes the scalar product. Note that if  $F^*$  is the solution of the optimization problem in (68), then  $C$  also is achieved by  $F^*$  in (69) and

$$\langle g(F^*), z_0^* \rangle = 0. \quad (70)$$

Now we need to verify that the Slater's condition holds, i.e. there exist an interior point  $F \in \Omega$  such that all the constraints hold with a strict inequality, i.e.  $g_i(F) < 0, i = 1, 2$ . Let  $x_1$  satisfies  $|x_1| < 0$  and  $\mathcal{E}(\hat{x}_1) > E_{\text{req}}$  and consider  $F_1$  the step function at  $x_1$ , then the following holds

$$g_1(F_1) = x_1^2 - P < 0, \quad (71)$$

$$g_2(F_1) = -\mathcal{E}(\hat{x}_1) + E_{\text{req}} < 0. \quad (72)$$

Hence the conditions of the Lagrangian theorem are satisfied, which completes the proof.

#### APPENDIX C PROOF OF THEOREM 2

The proof relies on the basic optimization Theorem [1], that determine a necessary and sufficient condition for the optimal solution of the following optimization problem. Let  $f$  be a continuous, and weakly-differentiable, and strictly convex map from a compact and convex space  $\Omega$  to  $\mathbb{R}$ . Define also

$$C \triangleq \sup_{x \in \Omega} f(x), \quad (73)$$

then the following claims hold,

- 1)  $C = \max f(x) = f(x_0)$  for some unique  $x_0 \in \Omega$ .
- 2) A necessary and sufficient condition for  $f(x_0) = C$  is  $f'_{x_0}(x) \leq 0$ .

Since  $\Omega$  is a convex, and  $J(F) \triangleq I(F) - \lambda_1 g_1(F) - \lambda_2 g_2(F)$  is weakly differentiable, then by applying the optimization theorem,  $J'_{F^*}(F) \leq 0$  is a necessary and sufficient condition for  $J(F)$  to achieve its maximum on  $F^*$ , where  $J'_{F^*}(F)$  is defined as follows

$$J'_{F^*}(F) = \lim_{\theta \rightarrow 0} \frac{J((1-\theta)F^* + \theta F) - J(F^*)}{\theta}. \quad (74)$$

It remains to prove that  $F \rightarrow J(F)$  is weakly differentiable and determine its first derivative. Define

$$F_\theta = (1-\theta)F^* + \theta F. \quad (75)$$

Then

$$\begin{aligned} I(F_\theta) - I(F^*) &= \int \int p(y|x) \log \frac{p(y; F^*)}{p(y; F_\theta)} dy dF^*(x) \\ &+ \theta \int \int i(x; F_\theta) dF(x) - \theta \int \int i(x; F_\theta) dF^*(x). \end{aligned} \quad (76)$$

Since

$$p(y; (1-\theta)F^* + \theta F) = (1-\theta)p(y; F^*) + \theta p(y; F), \quad (77)$$

the following expression holds

$$I'_{F^*}(F) = \lim_{\theta \rightarrow 0} \frac{I(F_\theta) - I(F^*)}{\theta} = \int i(x; F^*) dF(x) - I(F^*). \quad (78)$$

It has been shown that for the linear constraints,

$$g'_{i,F^*}(F) = g_i(F) - g_i(F^*), \quad (79)$$

and from the complementary slackness conditions,

$$g_i(F^*) = 0, \quad i \in \{1, 2\}. \quad (80)$$

Hence the condition  $J'_{F^*}(F) \leq 0$  implies

$$\int i(x; F^*) - C - \lambda_1 g_1(F) - \lambda_2 g_2(F) dF(x) \leq C - \lambda_1 P + \lambda_2 E_{\text{req}}, \quad (81)$$

which completes the proof.

#### APPENDIX D PROOF OF COROLLARY 2

Let  $\text{Supp}(F^*)$  be the points of support of a distribution function  $F^*$  and define

$$A_1(x) = x^2, \quad (82)$$

$$A_2(x) = -\mathcal{E}(\hat{x}), \quad (83)$$

$$a_1 = P, \quad (84)$$

$$a_2 = -E_{\text{req}}. \quad (85)$$

Thus we can write the inequality in Theorem 2 as following

$$\int \left( i(x; F^*) - \sum_{i=1}^2 \lambda_i A_i(x) \right) dF(x) \leq C - \sum_{i=1}^2 a_i^2. \quad (86)$$

Now, we need to prove that (86) is satisfied if and only if

$$i(x; F^*) \leq C + \sum_{i=1}^2 \lambda_i (A_i(x) - a_i), \quad \text{for } x \in [0, A], \quad (87)$$

and

$$i(x; F^*) = C + \sum_{i=1}^2 \lambda_i (A_i(x) - a_i), \quad \text{for } x \in \text{Supp}(F^*). \quad (88)$$

Note that if (87) and (88) hold then (86) is satisfied immediately. The second part of the proof is to prove (by contradiction) that if (86) holds then (87) and (88) are satisfied. First, we assume that if the inequality in (87) is false then there exist  $\tilde{x}$ , such that

$$i(\tilde{x}; F^*) > C + \sum_{i=1}^2 \lambda_i (A_i(\tilde{x}) - a_i), \quad \text{for all } x \in \text{Supp}(F^*). \quad (89)$$

Since (89) holds  $\forall F \in \Omega$ , then by choosing a particular distribution as the step function at  $\tilde{x}$ , the following holds

$$i(\tilde{x}; F^*) > C + \sum_{i=1}^2 \lambda_i (A_i(\tilde{x}) - a_i), \quad (90)$$

which is a contradiction to the inequality in (86). Now assume (87) holds but not (88), then there exist  $\tilde{x} \in \text{Supp}(F^*)$  such that

$$i(\tilde{x}; F^*) > C + \sum_{i=1}^2 \lambda_i (A_i(\tilde{x}) - a_i). \quad (91)$$

By using the continuity of the functions in (91), then there exist a set  $E'$  (neighborhood of  $\tilde{x}$ ) with a non zero measure, i.e.,  $\int_{E'} dF^*(x) = \delta > 0$ , such that (91) holds. Hence,

$$\begin{aligned} C - \sum_{i=1}^2 \lambda_i a_i &= I(F^*) - \sum_{i=1}^2 \lambda_i \int A_i(x) dF^*(x) \\ &= \int \left( i(x; F^*) - \sum_{i=1}^2 \lambda_i A_i(x) \right) dF^*(x) \\ &= \int_{E'} \left( i(x; F^*) - \sum_{i=1}^2 \lambda_i A_i(x) \right) dF^*(x) \\ &\quad + \int_{E-E'} \left( i(x; F^*) - \sum_{i=1}^2 \lambda_i A_i(x) \right) dF^*(x) \\ &< \delta \left( C - \sum_{i=1}^2 \lambda_i a_i \right) + (1 - \delta) \left( C - \sum_{i=1}^2 \lambda_i a_i \right) \\ &< C - \sum_{i=1}^2 \lambda_i a_i, \end{aligned} \quad (92)$$

which is a contradiction. This completes the proof.

#### APPENDIX E PROOF OF THEOREM 3

Assuming that  $S^*$  is not discrete, and by using the fact that the interval  $\left[\frac{1}{1+A^2}, 1\right]$  is compact, hence the set  $\text{Supp}(S^*)$  has an accumulation point by the Bolzano-Weierstrass theorem [30] in  $\left[\frac{1}{1+A^2}, 1\right]$ . Denote by  $T : z \rightarrow T(z)$  the following function

$$\begin{aligned} g(z) &= \lambda_1 \left( \frac{1}{z} - 1 - a \right) - \lambda_2 \left( I_0(\sqrt{2}Bh_2 \left( \sqrt{\frac{1}{z} - 1} \right) - b \right) \\ &\quad + C - \log z + 1 + \int_0^\infty z e^{-zy} \log p(y) dy, \quad z \in \mathcal{D}, \end{aligned}$$

with  $\mathcal{D}$  defined by  $\Re(z) > 0$ . By extending the necessary and sufficient conditions of Proposition 1 to the complex domain, we have

$$T(z) = 0, \quad z \in \text{Supp}(S^*). \quad (93)$$

Recall that the support of  $S^*$  has an accumulation point and the function  $g(z)$  is analytic over the domain  $\mathcal{D}$ , hence by applying the identity theorem [1], we have

$$T(z) = 0, \quad z \in \mathcal{D}. \quad (94)$$

By using (94), we have

$$\begin{aligned} \int_0^\infty s e^{-sy} \log p(y) dy &= -\frac{1}{s} \left[ \lambda_1 \left( \frac{1}{s} - 1 - a \right) \right. \\ &\quad \left. - \lambda_2 \left( I_0(\sqrt{2}Bh_2 \left( \sqrt{\frac{1}{s} - 1} \right) - b \right) + C - \log s + 1 \right], \\ \forall s &\in \left[ \frac{1}{1+A^2}, 1 \right]. \end{aligned} \quad (95)$$

The left hand side in (95) is the unilateral Laplace transform of the function  $\log p(y)$ , while the right-hand side (without the Bessel function [31]) can be recognized as the Laplace transform of

$$-\lambda_1 y + [\lambda_1(1+P) - C - 1 - C_E] - \log y, \quad (96)$$

where  $C_E$  is Euler's constant. (96) follows from the following unilateral Laplace inverse transform as

$$\mathcal{L}^{-1} \left\{ -\lambda_1 \frac{1}{s^2} \right\} (y) = -\lambda_1 y, \quad (97)$$

$$\mathcal{L}^{-1} \left\{ -\frac{1}{s} (-1 - P + C + 1) \right\} (y) = \lambda_1(1+P) - C - 1, \quad (98)$$

$$\mathcal{L}^{-1} \left\{ \frac{1}{s} \log s \right\} (y) = -\log y - C_E, \quad (99)$$

where (99) follows from the fact that  $\mathcal{L} \{ \log(y) \} (s) = -\frac{C_E + \log s}{s}$  [?]. The modified Bessel function is given by

$$\begin{aligned} I_0 \left( \sqrt{2} B h_2 \left( \sqrt{\frac{1}{s} - 1} \right) \right) &= \sum_{n=0}^{\infty} a_n \left( \frac{1}{s} - 1 \right)^n \\ &= \sum_{n=0}^{\infty} a_n \sum_{k=0}^n \binom{n}{k} \frac{1}{s^k} (-1)^{n-k}, \end{aligned}$$

with  $a_n = \frac{(B h_2 / \sqrt{2})^{2n}}{n!^2}$ . By using the fact that [?]

$$\mathcal{L}^{-1} \left\{ \frac{1}{s^k} \right\} (y) = \frac{y^{n-1}}{n!}, \quad (100)$$

and by taking into account the uniqueness of the Laplace transform for continuous functions with bounded variation, the following expression holds

$$\begin{aligned} p(y) &= K \frac{\exp(-\lambda_1 y)}{y} \\ &\times \exp \left( \lambda_2 \sum_{n=0}^{\infty} a_n \sum_{k=0}^n \binom{n}{k} \frac{y^k}{(k+1)!} (-1)^{n-k} \right). \end{aligned} \quad (101)$$

For every  $\lambda_1 > 0$  and  $\lambda_2 > 0$ , we have

$$\int_0^{\infty} p(y) dy > \infty, \quad (102)$$

hence,  $p(y)$  cannot be a probability distribution, and  $\text{Supp}(S^*)$  can not have an accumulation point; which means that the optimal input distribution is discrete. It is worth noting that the proof of the discreteness does not take into account the HPA non-linearity, since the consideration of the HPA leads to a complicated expressions for the unilateral Laplace transform in (95). However, the discreteness property of the optimal input distribution holds for the HPA case as well. Specifically, by proving that the optimal input  $X^*$  is a discrete random variable, it is equivalent to prove that  $\hat{X}^* = \frac{X}{\left[1 + \left(\frac{X}{A_s}\right)^{2\beta}\right]^{\frac{1}{2\beta}}}$  is discrete.

## REFERENCES

- [1] J. G. Smith, "The information capacity of amplitude- and variance-constrained," *Inf. Control*, vol. 18, pp. 203–219, 1971.
- [2] S. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication: opportunities and challenges," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 117–125, Apr. 2015.
- [3] N. Tesla, "Apparatus for transmitting electrical energy," New York, NY, Dec. 1914.
- [4] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, "Simultaneous wireless information and power transfer in modern communication systems," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 104–110, Nov. 2014.
- [5] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, Canada, 2008, pp. 1612–1616.
- [6] P. Grover and A. Sahai, "Shannon meets tesla: Wireless information and power transfer," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, Texas, June. 2010, pp. 2363–2367.
- [7] S. Belhadj Amor, S. M. Perlaza, I. Krikidis, and H. V. Poor, "Feedback enhances simultaneous wireless information and energy transmission in multiple access channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 5244–5265, Aug. 2017.
- [8] N. Khalfet and S. M. Perlaza, "Simultaneous Information and Energy Transmission in the Two-User Gaussian Interference Channel," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 156–170, Jan. 2019.
- [9] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a two-user MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4210–4221, Aug. 2013.
- [10] T. Tu Lam, M. Di Renzo, and J. P. Coon, "System-level analysis of SWIPT mimo cellular networks," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 2011–2014, 2016.
- [11] S. M. Belhadj Amor and S. M. Perlaza, "Fundamental limits of simultaneous energy and information transmission," in *Proc. 23rd Int. Symp. Telecomm.*, Thessaloniki, Greece, May 2016.
- [12] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May. 2013.
- [13] B. Clerckx and E. Bayguzina, "Waveform design for wireless power transfer," *IEEE Trans. Sign. Proc.*, vol. 64, no. 23, pp. 6313–6328, 2016.
- [14] M. Varasteh, B. Rassouli, and B. Clerckx, "On capacity-achieving distributions for complex AWGN channels under nonlinear power constraints and their applications to SWIPT," *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6488–6508, Oct. 2020.
- [15] R. Morsi, V. Jamali, D. W. K. Ng, and R. Schober, "On the capacity of SWIPT systems with a nonlinear energy harvesting circuit," in *Proc. IEEE Int. Conf. Commun.*, Kansas City, USA, May. 2018, pp. 1–7.
- [16] A. Collado and A. Georgiadis, "Optimal waveforms for efficient wireless power transmission," *IEEE Micro Wireless Comp Lett*, vol. 24, no. 5, pp. 354–356, 2014.
- [17] D. I. Kim, J. H. Moon, and J. J. Park, "New swipt using papr: How it works," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 672–675, 2016.
- [18] J. Qi and S. Aissa, "Analysis and compensation of power amplifier nonlinearity in MIMO transmit diversity systems," *IEEE Trans. Vehic. Techn.*, vol. 59, no. 6, pp. 2921–2931, 2010.
- [19] P. Mukherjee, S. Lajnef, and I. Krikidis, "MIMO SWIPT systems with power amplifier nonlinearities and memory effects," *IEEE Wirless. Comm. Lett.*, pp. 1–1, 2020.
- [20] I. Krikidis, "Information-energy capacity region for SWIPT systems with power amplifier nonlinearity," in *Proc. IEEE Int. Symp. Inf. Theory*, LA, California, USA, 2020, pp. 3067–3072.
- [21] O. Ozel and S. Ulukus, "Achieving AWGN capacity under stochastic energy harvesting," *IEEE Trans Inf Theory*, vol. 58, no. 10, pp. 6471–6483, 2012.
- [22] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.
- [23] O. Ozel and S. Ulukus, "On the capacity region of the gaussian MAC with batteryless energy harvesting transmitters," in *proc IEEE Global Comm. Conf.*, 2012, pp. 2385–2390.
- [24] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290–1301, May. 2001.
- [25] R. Morsi, V. Jamali, A. Hagelauer, D. W. K. Ng, and R. Schober, "Conditional capacity and transmit signal design for swipt systems with multiple nonlinear energy harvesting receivers," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 582–601, 2020.

- [26] O. Ozel and S. Ulukus, "AWGN channel under time-varying amplitude constraints with causal information at the transmitter," in *proc Conf Asil. Conf. on Sign., Syst. and Comp.*, Pacific Grove, CA, 2011, pp. 373–377.
- [27] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx>, Mar. 2014.
- [28] N. Sharma and S. Shamai, "Transition points in the capacity-achieving distribution for the peak power limited awgn and free-space optical intensity channels," *Probl. Inf. Transm.*, vol. 46, pp. 283–299, 2010.
- [29] N. Sharma and S. Shamai, "Characterizing the discrete capacity achieving distribution with peak power constraint at the transition points," in *Proc Int. Symp. Inf. Theory and Its Applications*, Auckland, New Zealand, Dec 2008, pp. 1–6.
- [30] W. Feller, *An Introduction to Probability Theory and Its Application*, 2nd ed. New York, NY: John Wiley and Sons, 1971, vol. 2.
- [31] F. Bowman, *Introduction to Bessel Function*, 2nd ed. New York, NY: Dover Publications, 1958.



**Nizar Khalfet** received the BSc in electrical engineering from Sup'Com, Tunis, Tunisia, and the MSc degree from CentraleSupélec, Paris, France in 2015, and the Ph.D. degree in electrical engineering from the Institut National de Recherche en Informatique et en Automatique (INRIA), France. He is currently a Researcher with the IRIDA Research Centre for Communication Technologies, University of Cyprus since 2020. His research interests include topics in the intersection of information theory, intelligent reflected surfaces and communications theory.



**Ioannis Krikidis** (Fellow, IEEE) received the Diploma degree in computer engineering from the Computer Engineering and Informatics Department (CEID), University of Patras, Greece, in 2000, and the M.Sc. and Ph.D. degrees in electrical engineering from the Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, in 2001 and 2005, respectively. From 2006 to 2007, he worked as a Post-Doctoral Researcher with ENST and a Research Fellow with the School of Engineering and Electronics, The University of Edinburgh, Edinburgh, U.K., from 2007 to 2010. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus. His current research interests include wireless communications, cooperative networks, 5G/B5G communication systems, wireless powered communications, and intelligent reflected surfaces. He was a recipient of the 2013 Young Researcher Award from the Research Promotion Foundation, Cyprus, the 2016 IEEEComSoc Best Young Professional Award in Academia, and the 2019 IEEE SIGNAL PROCESSING LETTERS Best Paper Award. He has been recognized by the Web of Science as a Highly Cited Researcher for 2017–2021. He has received the prestigious ERC consolidator grant. He serves as an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, and IEEE WIRELESS COMMUNICATIONS LETTERS.

Edinburgh, U.K., from 2007 to 2010. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus. His current research interests include wireless communications, cooperative networks, 5G/B5G communication systems, wireless powered communications, and intelligent reflected surfaces. He was a recipient of the 2013 Young Researcher Award from the Research Promotion Foundation, Cyprus, the 2016 IEEEComSoc Best Young Professional Award in Academia, and the 2019 IEEE SIGNAL PROCESSING LETTERS Best Paper Award. He has been recognized by the Web of Science as a Highly Cited Researcher for 2017–2021. He has received the prestigious ERC consolidator grant. He serves as an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, and IEEE WIRELESS COMMUNICATIONS LETTERS.