Wireless Information and Power Transfer over an AWGN channel: Nonlinearity and Asymmetric Gaussian Signalling

Morteza Varasteh, Borzoo Rassouli, Bruno Clerckx

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

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

Abstract

Simultaneous transmission of information and power over a point-to-point flat-fading complex Additive White Gaussian Noise (AWGN) channel is studied. In contrast with the literature that relies on an inaccurate linear model of the energy harvester, an experimentally-validated nonlinear model is considered. A general form of the delivered Direct Current (DC) power in terms of system baseband parameters is derived, which demonstrates the dependency of the delivered DC power on higher order statistics of the channel input distribution. The optimization problem of maximizing Rate-Power (R-P) region is studied. Assuming that the Channel gain is available at both the receiver and the transmitter, and constraining to independent and identically distributed (i.i.d.) channel inputs determined only by their first and second moment statistics, an inner bound for the general problem is obtained. Notably, as a consequence of the harvester nonlinearity, the studied inner bound exhibits a tradeoff between the delivered power and the rate of received information. It is shown that the tradeoff-characterizing input distribution is with mean zero and with asymmetric power allocations to the real and imaginary dimensions.

I. INTRODUCTION

Radio-Frequency (RF) waves can be utilized for transmission of both information and power simultaneously. 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 (DMC). He showed the existence of tradeoff between the information rate and the delivered power for some channels, such as, point-to-point binary channels and amplitude constraint Gaussian channels. Recent results in the literature have also revealed that in many scenarios, there is a tradeoff between information

rate and delivered power. Just to name a few, frequency-selective channel [2], MIMO broadcasting [3], interference channel [4], [5], relaying [6], [7].

One of the major efforts in a Simultaneous Wireless Information and Power Transfer (SWIPT) architecture is to increase the Direct-Current (DC) power at the output of the harvester without increasing transmit power. The harvester, known as rectenna, is composed of a rectifier[1] followed by a low-pass filter. In [8], [9], it is shown that 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. In [9], an analytical model for rectenna's output is introduced via the Taylor expansion of the diode characteristic function. As one of the main conclusions, it is shown that the rectifier's nonlinearity is key to design efficient wireless powered systems.

The design of an efficient SWIPT architecture fundamentally relies on designing an efficient Wireless Power Transfer (WPT) structure as an important building block of SWIPT. The SWIPT literature has so far focused on the linear model of the rectifier, e.g., [2]–[7], whereas, it is expected that considering nonlinearity effect changes the SWIPT design, signalling and architecture significantly. Indeed, in [10], [11], the design of SWIPT waveforms is studied accounting for rectenna's nonlinearity with a power splitter at the receiver. It is shown that superposing deterministic multisines (for power transfer purposes) with Orthogonal Frequency Division Multiplexing (OFDM) symbols modulated with Circularly Symmetric Complex Gaussian (CSCG) zero-mean inputs (for information purposes) enlarges the Rate-Power (R-P) region, compared to merely zero-mean inputs. This highlights the potential and benefits of departing from conventional CSCG inputs in SWIPT.

Leveraging the aforementioned observations, we provide a step closer at identifying the fundamental limits of SWIPT accounting for the nonlinearity of rectenna. In this paper, we study a flat-fading Additive White Gaussian Noise (AWGN) channel for SWIPT. Taking the advantage of the approximation for rectenna's nonlinear output introduced in [9], we obtain the general form of the delivered power in terms of system baseband parameters. Assuming that the receiver jointly extracts information and harvests power from the received RF signal,[2] it is shown that the delivered power at the receiver is dependent on the

---

[1]In the literature, the rectifier is usually considered as a diode, which is the main source of nonlinearity induced in the system.

[2]We note that, leveraging the results in thermodynamics of computing, it is demonstrated that energy need not be dissipated in the decoding process. This is due to the reason that to perform a mathematical work, energy is not required [12, Ch. 5]. In particular, decoders that are reversible computational devices would not dissipate any energy [13] and electronic circuits that are almost thermodynamically reversible have been built [14]. Motivated by this, we also assume that at the receiver, the decoder is able to jointly harvest power and extract information from the received RF signal.

first to fourth moment statistics of the channel input distribution. Considering the optimization problem of maximizing R-P region, we obtain an achievable scheme as an inner bound for the general problem. The scheme is based on constraining the channel inputs to independent and identically distributed (i.i.d.) distributions that are determined by their first and second moment statistics. For the studied inner bound, we show that there is a tradeoff between the delivered power at the receiver and the rate of the received information. This result is highlighted in contrast to the scenario, in which a linear model is considered for the power harvester at the receiver. It can be easily verified that under an assumption of linear model for the power harvester, the goal of maximum rate and maximum energy are aligned in the flat-fading channel. Additionally, we show that the maximum rate-power (for the studied inner bound) is achieved when the channel input distributions is Gaussian with mean zero, however, with different (asymmetric) power allocations to the real and imaginary dimensions.

Organization: In Section II, we introduce the system model. In Section III, the delivered power at the receiver is obtained in terms of system baseband parameters accounting the approximation for nonlinearity of rectenna. In Section IV, we introduce the problem considered in this paper, and accordingly, in Section V, we obtain an achievable scheme as an inner bound for the general optimization problem. In Section VI, we conclude the paper.

Notation: Throughout this paper, the standard CSCG distribution is denoted by $\mathcal{CN}(0, 1)$. Complex conjugate of a complex number c is denoted by c^*. For a random process $X(t)$, corresponding random variable at time index n is represented by X_n. The operators $\mathbb{E}[\cdot]$ and $\mathcal{E}[\cdot]$ denote the expectation over statistical randomness and the average over time, respectively. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are real and imaginary operators, respectively. We use the notations $\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$ and $s_l = \mathrm{sinc}(l + 1/2)$ for integer l. We also define δ_l as

$$\delta_l = \begin{cases} 1 & l = 0 \\ 0 & l \neq 0 \end{cases}.$$ (1)

II. SYSTEM MODEL

Considering a point-to-point flat-fading AWGN channel, in the following, we explain the operation of the transmitter and the receiver.

A. Transmitter

At the transmitter, the signal $X(t)$ is produced as

$$X(t) = \sum_n X_n \mathrm{sinc}(f_w t - n),$$ (2)

where X_n is an information-power symbol at time index n, modelled as a random variable, which is produced in an i.i.d. fashion and $X(t)$ is with bandwidth $[-f_w/2, f_w/2]$. Next, the signal $X(t)$ is upconverted to the carrier frequency f_c and is sent over the channel.

B. Receiver

The filtered received RF waveform at the receiver is modelled as

$$Y_{\text{rf}}(t) = \sqrt{2}\Re\left\{Y(t)e^{j2\pi f_c t}\right\}, \tag{3}$$

where $Y(t)$ is the baseband equivalent of the channel output with bandwidth $[-f_w/2, f_w/2]$. We assume that $f_c > 2f_w$.

Power: At the receiver, the power of the RF signal $Y_{\text{rf}}(t)$ is delivered through the rectenna. In the following, we leverage the approximation for rectenna's output introduced in [9][3]. Accordingly, the delivered power (denoted by P_{del}) is modelled as[4]

$$P_{\text{del}} = \mathbb{E}\mathcal{E}[k_2 Y_{\text{rf}}(t)^2 + k_4 Y_{\text{rf}}(t)^4], \tag{4}$$

where k_2 and k_4 are constants. Note that, in the linear model for the delivered power P_{del}, in (4), we have only the second moment of the received RF signal $Y_{\text{rf}}(t)$. Validating through circuit simulations in [9], it is shown that the linear model is inaccurate and inefficient from a signal design perspective.

Information: The signal $Y_{\text{rf}}(t)$ is downconverted producing the baseband signal $Y(t)$ given as [5]

$$Y(t) = \sum_i a_i^b(t)X(t - \tau_i(t)) + W(t). \tag{5}$$

Next, $Y(t)$ is sampled with a sampling frequency f_w producing $Y_m = Y(m/f_w)$ given as

$$Y_m = X_m \sum_i a_i^b(m/f_w) + W_m, \tag{6}$$

where in (6), we used $\tau_i(m/f_w)f_w \approx 0$ because the channel is flat-fading. W_m and X_m represent samples of the additive noise $W(t)$ and the signal $X(t)$ at time $t = m/f_w$, respectively.

We model W_m as an i.i.d. and CSCG random variable with variance σ_w^2, i.e., $W_m \sim \mathcal{CN}(0, \sigma_w^2)$. We assume that both the transmitter and the receiver know the Channel gain, namely, $h(t) = \sum_i a_i^b(t)$ at times

---

[3]According to [9], due to the presence of a diode in rectenna's structure, its output current is an exponential function, which is approximated by expanding its Taylor series. The approximation used here, is the fourth moment truncation of Taylor series, in which the first and third moments are zero with respect to the time averaging.

[4]According to [9], rectenna's output in (4) is in the form of current with unit Ampere. However, since power is proportional to current, with abuse of notation, we refer to the term in (4) as power.

[5]We model the baseband equivalent channel impulse response as $H(\tau, t) = \sum_i a_i^b(t)\delta(\tau - \tau_i(t)) + W(t)$ where $\alpha_i^b(t), \tau_i(t)$ are the channel coefficient and delay of path i.

$t = m/2f_w$ for integer m, which is assumed to be fixed over all the transmissions. Throughout the paper, since the transmitted symbols X_m and the noise W_m are i.i.d., we drop the index m for X_m, W_m. We also define $h = \sum_i a_i^b((2m)/(2f_w))$ and $\tilde{h} = \sum_i a_i^b((2m+1)/(2f_w))$. Note that h and $\tilde{h}$ are assumed to be fixed, however, we assume they are not necessarily equal. Therefore, (6) reads

$$Y = hX + W. \tag{7}$$

Note that in (7), only even samples of the channel, i.e., h are involved.

III. DELIVERED POWER

In this section, we study the power delivered at the receiver. Note that most of the communication processes, such as, coding/ decoding, modulation/ demodulation, etc, is done at the baseband. Therefore, from a communication system design point of view, it is most preferable to have baseband equivalent presentation of the system. Henceforth, in the following Proposition, we derive the delivered power P_{del} at the receiver in terms of system baseband parameters.

Proposition 1. *Assuming the channel input distributions are i.i.d., the delivered power P_{del} at the receiver, can be expressed in terms of system baseband parameters as*

$$P_{del} = \alpha Q + \tilde{\alpha}\tilde{Q} + (\beta + \tilde{\beta})P + \gamma, \tag{8}$$

where $\tilde{Q}$ is given by

$$\tilde{Q} = \frac{1}{3}\big(Q_r + Q_i + 2(\mu_r T_r + \mu_i T_i)$$
$$+ 6P_r P_i + 6P_r(P_r - \mu_r^2) + 6P_i(P_i - \mu_i^2)\big), \tag{9}$$

where the parameters α, $\tilde{\alpha}$, β, $\tilde{\beta}$ and γ are given as

$$\alpha = \frac{3k_4}{4f_w}|h|^4, \tag{10}$$

$$\tilde{\alpha} = \frac{3k_4}{4f_w}|\tilde{h}|^4, \tag{11}$$

$$\beta = \frac{1}{f_w}\left(k_2 + 6k_4\sigma_w^2\right)|h|^2, \tag{12}$$

$$\tilde{\beta} = \frac{1}{f_w}\left(k_2 + 6k_4\sigma_w^2\right)|\tilde{h}|^2, \tag{13}$$

$$\gamma = \frac{1}{f_w}(k_2\sigma_w^2 + 3k_4\sigma_w^4), \tag{14}$$

and $Q = \mathbb{E}[|X|^4]$, $T = \mathbb{E}[|X|^3]$, $P = \mathbb{E}[|X|^2]$, $\mu = \mathbb{E}[X]$. Similarly, $Q_r = \mathbb{E}[\Re\{X\}^4]$, $T_r = \mathbb{E}[\Re\{X\}^3]$, $P_r = \mathbb{E}[\Re\{X\}^2]$, $\mu_r = \mathbb{E}[\Re\{X\}]$ and $Q_i = \mathbb{E}[\Im\{X\}^4]$, $T_i = \mathbb{E}[\Im\{X\}^3]$, $P_i = \mathbb{E}[\Im\{X\}^2]$, $\mu_i = \mathbb{E}[\Im\{X\}]$.

Proof: See Appendix A.

Remark 1. *We note that obtaining a closed form expression for the delivered power P_{del} at the receiver when the channel inputs are not i.i.d. is cumbersome. This is due to the fact that the fourth moment of the received signal $Y_{rf}(t)$ creates dependencies of the statistics of the present channel input on the statistics of the channel inputs on other time indices (see e.g., eq. (54) and eq. (50) in Appendix A).*

IV. PROBLEM STATEMENT

We aim at maximizing the rate of the received information, as well as the amount of power delivered at the receiver. Accordingly, the optimization problem we consider, is the maximization of mutual information between the channel input X and the channel output Y under a given power constraint at the transmitter and a minimum delivered power constraint at the receiver. Hence, for the optimization problem, we have

$$
\sup_{p_X(x)} \quad I(X;Y)
$$
$$
\text{s.t.} \quad \begin{cases} P \leq P_a \\ P_{\text{del}} \geq P_d \end{cases}, \tag{15}
$$

where $\sup$ is taken over all input distributions $p_X(x)$ satisfying the constraints in (15). P_a is the available power budget at the transmitter and P_d is the minimum amount of power that is to be delivered to the receiver.

Remark 2. *We note that, for the problem in (15), if the second constraint (the minimum delivered power at the receiver) is represented via a linear model, i.e., $\mathbb{E}[|Y|^2] \geq P_d$, the maximum is achieved using a CSCG input distribution. It can also be verified easily that there is no tradeoff between the received information rate and delivered power at the receiver.*

V. MAIN RESULT

In this section, we obtain an inner bound for the problem in (15) by constraining the input distributions to those that are determined by their first and second moment statistics[6]. We show that for the considered scenario, there is a tradeoff between the rate of the transmitted information, namely $I(X;Y)$ and delivered power P_{dc} at the receiver and accordingly, we characterize the tradeoff.

Proposition 2. *When a channel input distribution $p_X(x)$ is completely determined by its first and second moment statistics, the supremum in (15) is achieved by a zero mean Gaussian distribution as the channel*

---

[6]This assumption is justified due to the fact that in practice, most of the modulation schemes are i.i.d. and are fully characterized by the knowledge of the first and second moment statistics only.

input, i.e., $\Re\{X\} \sim \mathcal{N}(0, P_r)$, and $\Im\{X\} \sim \mathcal{N}(0, P_i)$, where $P = P_r + P_i = P_a$. Furthermore, let $P_{dc,max} = 3(\alpha + \tilde{\alpha})P_a{}^2 + (\beta + \tilde{\beta})P_a + \gamma$ and $P_{dc,min} = 2(\alpha + \tilde{\alpha})P_a{}^2 + (\beta + \tilde{\beta})P_a + \gamma$ be the maximum and minimum delivered power at the receiver, respectively. For $P_d = P_{dc,max}$, the maximum in (15) is attained by $P_i = 0, P_r = P_a$ or $P_i = P_a, P_r = 0$. For $P_d = P_{dc,min}$, the maximum in (15) is attained by $P_i = P_a/2, P_r = P_a/2$. For $P_{dc,min} < P_d < P_{dc,max}$, the optimal power allocation that attains the maximum rate is given by P_i^* and $P_r^* = P_a - P_i^*$, where P_i^* is chosen such that the following equation is satisfied

$$\alpha Q + \tilde{\alpha}\tilde{Q} + (\beta + \tilde{\beta})P_a + \gamma = P_d. \tag{16}$$

For $P_d < P_{dc,min}$, the optimal power allocation is attained by $P_i^* = P_r^* = P_a/2$ and the delivered power is still $P_{dc,min}$.

Proof: See Appendix B.

Note that in (16), for a complex zero-mean Gaussian distributed channel input with P_r and P_i as the variances of real and imaginary dimensions, respectively, we have

$$Q = \tilde{Q} = 3(P_i^2 + P_r^2) + 2P_i P_r. \tag{17}$$

Remark 3. *From (8), it is seen that the delivered power P_{del} at the receiver depends on the second moment statistics P_r, P_i, as well as the fourth moment statistics Q_r, Q_i of the channel input X. This is due to the presence of the fourth moment of the received signal Y_{rf} in modelling the rectenna's output. Accordingly, the rate is minimized (corresponding to $P_{dc,max}$) when the available power at the transmitter is fully allocated to one of the real or imaginary dimensions. This is because allocating power to one dimension, leads to a higher fourth order moment. On the other hand, the maximum rate is achieved (corresponding to $P_{dc,min}$) when the available power is equally distributed between the real and the imaginary dimensions. This is elaborated in Figure 1.*

Remark 4. *As mentioned earlier, maximization of both the delivered power P_{del} and the rate $I(X;Y)$ are aligned under the linear modelling for the delivered power. Therefore, regardless of the channel condition, the best power allocation is always $P_i = P_r = P/2$. However, accounting the fourth moment of the received signal Y_{rf} in (4), the receiver chooses the proper power allocation such that the constraints in (15) are satisfied (if $P_d \leq P_{dc,max}$). We also note that since we have assumed the channel is fixed for the whole transmission, therefore, the transmitter keeps using the same power allocation for the whole transmission.*

It is also noted that at each time index the rate of the information $I(X;Y)$ is affected only through h, whereas, the delivered power P_{del} is affected through both h and $\tilde{h}$. This is illustrated in Figure 1 by representing three different regions by varying the channel coefficients.
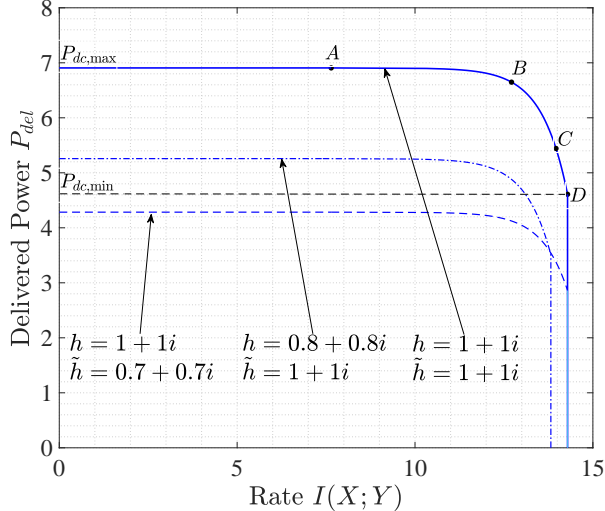
Figure 1. Delivered power P_{del} and achievable information rate $I(X; Y)$ for the problem in (15) for i.i.d. channel inputs that are determined by their first and second moment statistics ($P_a = 1$, $\sigma_w^2 = 10^{-4}$, $f_w = 1$, $k_2 = 0.17$, $k_4 = 19.145$). The points A, B, C, D correspond to $(P_r, P_i) = (0, 1)$, $= (0.03, 0.97)$, $= (0.2, 0.8)$, $= (0.5, 0.5)$, respectively, where the values for P_r and P_i can be interchanged without affecting the resulting points. The values for k_2, k_4 are adapted from [10].

VI. CONCLUSION

In this paper, we studied SWIPT over a point-to-point complex AWGN channel in the presence of a nonlinear power harvester at the receiver. Assuming that the channel state information is available at both the transmitter and the receiver, we studied the problem of maximizing rate of the transmitted information as well as delivered power at the receiver. Assuming that the channel inputs are i.i.d. and are fully characterized by the knowledge of their first and second moment statistics, we derived an inner bound for the optimal R-P region. We showed that for the obtained inner bound, there is a trade off (due to the nonlinearity of the power harvester at the receiver) between the rate of transmitted information and delivered power. Accordingly, we characterized the inner bound, which demonstrates that the optimal channel input is still a zero mean Gaussian distribution, however, with asymmetric power allocations to the real and imaginary dimensions.

Among open problems that are left for future research, we mention here the optimal input distribution for the problem in (15). Another interesting problem is the extension of the problem studied in this paper to the frequency-selective AWGN channel.

A. PROOF OF PROPOSITION 1

The following series will be useful throughout the proof of the proposition 1.

Lemma 3. *Recalling that $s_l = sinc(l + 1/2)$ for integer l, we have the following series:*

$$S_0 \triangleq \sum_l s_l^2 = 1, \tag{18}$$

$$S_1 \triangleq \sum_l \sum_{k:k \neq l} s_l s_k = 0, \tag{19}$$

$$S_2 \triangleq \sum_l \sum_{k:k \neq l} \sum_{\substack{d:d \neq l \\ d \neq k}} \sum_{\substack{m:m \neq l \\ m \neq d \\ m \neq k}} s_l s_k s_d s_m = 0, \tag{20}$$

$$S_3 \triangleq \sum_l \sum_{k:k \neq l} s_l^2 s_k^2 = \frac{2}{3}, \tag{21}$$

$$S_4 \triangleq \sum_l \sum_{k:k \neq l} \sum_{\substack{d:d \neq l \\ d \neq k}} s_l^2 s_k s_d = -\frac{1}{3}, \tag{22}$$

$$S_5 \triangleq \sum_l s_l^4 = \frac{1}{3}, \tag{23}$$

$$S_6 \triangleq \sum_l \sum_{k:k \neq l} s_l^3 s_k = \frac{1}{2}. \tag{24}$$

Proof: See Appendix C.

Considering first the term $\mathbb{E}\mathcal{E}[Y_{\mathrm{rf}}(t)^2]$, we have

$$\mathbb{E}\mathcal{E}[Y_{\mathrm{rf}}(t)^2] = \frac{1}{2}\mathbb{E}\mathcal{E}\left[\left(Y(t)e^{jf_c t} + Y^*(t)e^{-jf_c t}\right)^2\right] \tag{25}$$

$$= \mathbb{E}\mathcal{E}\left[|Y(t)|^2\right] \tag{26}$$

$$= \mathbb{E}\mathcal{E}\left[\sum_{n,m} Y_n Y_m^* \mathrm{sinc}(f_w t - n)\mathrm{sinc}(f_w t - m)\right] \tag{27}$$

$$= \sum_{n,m} \mathbb{E}\left[Y_n Y_m^*\right] \mathcal{E}\left[\mathrm{sinc}(f_w t - n)\mathrm{sinc}(f_w t - m)\right] \tag{28}$$

$$= \lim_{T \to \infty} \frac{1}{f_w T} \sum_m \mathbb{E}\left[|Y_m|^2\right] \tag{29}$$

$$= |h|^2 P + \sigma_w^2, \tag{30}$$

where (26) is because we have $\mathcal{E}\{Y(t)^2 e^{2jf_c t}\} = 0$. (27) is due to the fact that the signal $Y(t)$ is bandlimited to f_w and we have

$$Y(t) = \sum_n Y_n \mathrm{sinc}(f_w t - n). \tag{31}$$

In (29), we used the equation

$$\mathcal{E}\left[\mathrm{sinc}(f_w t - n)\mathrm{sinc}(f_w t - m)\right] = \lim_{T \to \infty} \frac{1}{f_w T}\delta_{n-m}. \tag{32}$$

Considering the term $\mathbb{E}\mathcal{E}[Y_{\mathrm{rf}}(t)^4]$, we have

$$\mathbb{E}\mathcal{E}[Y_{\mathrm{rf}}(t)^4] = \frac{1}{4}\mathbb{E}\mathcal{E}\big[4|Y(t)|^4$$

$$+ (Y(t)^2 e^{j2f_c t} + Y^*(t)^2 e^{-j2f_c t})^2$$

$$+ 4|Y(t)|^2 (Y(t)^2 e^{j2f_c t} + Y^*(t)^2 e^{-j2f_c t})\big] \tag{33}$$

$$= \frac{3}{2}\mathbb{E}\mathcal{E}\big[|Y(t)|^4\big]. \tag{34}$$

Note that, the signal $|Y(t)|^2$ is real with bandwidth $[-f_w, f_w]$. Hence, it can be represented by its samples taken each $t = 1/2f_w$ seconds. Therefore, we have

$$|Y(t)|^2 = \sum_n Y_n^s \mathrm{sinc}(2f_w t - n), \tag{35}$$

where $Y_n^s \triangleq |Y(n/2f_w)|^2$. Accordingly, (34) reads as

$$\frac{3}{2}\mathbb{E}\mathcal{E}\big[|Y(t)|^4\big] = \frac{3}{2f_w}\sum_n \mathbb{E}[|Y_n^s|^2] \tag{36}$$

$$= \lim_{T \to \infty} \frac{3}{2Tf_w}\sum_k \mathbb{E}[|Y_{2k+1}^s|^2] + \frac{3}{2Tf_w}\sum_k \mathbb{E}[|Y_{2k}^s|^2]. \tag{37}$$

Note that $Y_{2k}^s = |Y(2k/2f_w)|^2 = |Y_k|^2$. Hence, $\mathbb{E}[|Y_{2k}^s|^2]$ in (37) reads

$$\mathbb{E}[|Y_{2k}^s|^2] = \mathbb{E}[|Y_k|^4] \tag{38}$$

$$= \mathbb{E}[((hX + W)(h^*X^* + W^*))^2] \tag{39}$$

$$= \mathbb{E}[|h|^4|X|^4 + |W|^4 + 2|h|^2|X|^2|W|^2$$

$$+ W^2 h^{*2} X^{*2} + W^{*2} h^2 X^2 + 2|h|^2|X|^2|W|^2$$

$$+ 2(|h|^2|X|^2 W h^* X^* + |h|^2|X|^2 W^* hX$$

$$+ |W|^2 W h^* X^* + |W|^2 W^* hX)] \tag{40}$$

$$= |h|^4 \mathbb{E}[|X|^4] + 2\sigma_w^4$$

$$+ 2\sigma_w^2 |h|^2 \mathbb{E}[|X|^2] + 2\sigma_w^2 |h|^2 \mathbb{E}[|X|^2] \tag{41}$$

$$= |h|^4 Q + 4\sigma_w^2 |h|^2 P + 2\sigma_w^4. \tag{42}$$

To calculate the term $\mathbb{E}[|Y_{2k+1}^s|^2]$ in (37), we note that the channel's baseband equivalent signal $Y(t)$ can be written as

$$Y(t) = \sum_n X_n \sum_i a_i^b(t)\mathrm{sinc}(f_w t - n) + W(t), \tag{43}$$

where we have neglected the term $f_w \tau_i$, since the channel is flat and we have $f_w \tau_i \approx 0$. Substituting $t = (2k+1)/f_w$ we have

$$\tilde{Y}_k \triangleq Y\left(\frac{2k+1}{2f_w}\right) \tag{44}$$

$$= \sum_n X_n \sum_i a_i^b \left(\frac{2k+1}{2f_w}\right) s_{k-n} + W\left(\frac{2k+1}{2f_w}\right) \tag{45}$$

$$= \sum_n X_n s_{k-n} \sum_i a_i^b \left(\frac{2k+1}{2f_w}\right) + W\left(\frac{2k+1}{2f_w}\right) \tag{46}$$

$$= \tilde{h}\tilde{X} + \tilde{W}. \tag{47}$$

where $\tilde{X} \triangleq \sum_n X_n s_{k-n}$ and $\tilde{W} \triangleq W((2k+1)/2f_w)$. Similarly to (42), we have

$$\mathbb{E}[|Y_{2k+1}^s|^2] = \mathbb{E}[|\tilde{Y}_k|^4] \tag{48}$$

$$= |\tilde{h}|^4 \tilde{Q} + 4\sigma_w^2 |\tilde{h}|^2 \tilde{P} + 2\sigma_w^4, \tag{49}$$

where $\tilde{Q} = \mathbb{E}[|\tilde{X}|^4]$, $\tilde{P} = \mathbb{E}[|\tilde{X}|^2]$. For $\tilde{P}$, we have

$$\tilde{P} = \mathbb{E}\left[\sum_{n,m} X_n X_m^* s_{k-n} s_{k-m}\right] \tag{50}$$

$$= \sum_{n,m:n=m} \mathbb{E}[|X_n|^2] s_{k-n}^2$$

$$+ \sum_{n,m:n\neq m} \mathbb{E}[X_n]\mathbb{E}[X_m^*] s_{k-n} s_{k-m} \tag{51}$$

$$= S_0 P + S_1 |\mu|^2 \tag{52}$$

$$= P, \tag{53}$$

where in (51) we used the assumption that X_n is i.i.d. with respect to different values of n. For $\tilde{Q}$, we have

$$\tilde{Q} = \mathbb{E}\left[\sum_{l,k,d,m} X_l X^* X_d^* X_m^* s_{n-l} s_{n-k} s_{n-d} s_{n-m}\right]. \tag{54}$$

Accounting for the different cases for the possible values of l, k, d, m, we have

- If all the indices l, k, d, m are with different values, we have

$$\tilde{Q} = |\mu|^4 S_2. \tag{55}$$

- If $(l = k, \ d \neq k, \ d = m)$ or $(l = d, \ k \neq d, \ k = m)$, we have

$$\tilde{Q} = P^2 S_3. \tag{56}$$

- If $(l = m, \ k \neq m, \ k = d)$, we have

$$\tilde{Q} = |\bar{P}|^2 S_3. \tag{57}$$

- If $(l = k, \ d \neq m, \ d \neq k, \ m \neq k)$ or $(l = d, \ k \neq m, \ k \neq d, \ m \neq d)$ or $(k = m, \ l \neq d, \ l \neq m, \ d \neq m)$ or $(d = m, \ l \neq k, \ l \neq m, \ k \neq m)$, we have

$$\tilde{Q} = P|\mu|^2 S_4. \tag{58}$$

- If $(l = m, \ k \neq d, \ k \neq m, \ d \neq m)$, we have

$$\tilde{Q} = \bar{P}\mu^{*2} S_4. \tag{59}$$

- If $(k = d, \ l \neq m, \ l \neq d, \ m \neq d)$, we have

$$\tilde{Q} = \bar{P}^* \mu^2 S_4. \tag{60}$$

- If $l = k = d = m$, we have

$$\tilde{Q} = Q S_5. \tag{61}$$

- If $l = k = d \neq m$ or $k = d = m \neq l$, we have

$$\tilde{Q} = \bar{T}^* \mu S_6. \tag{62}$$

- If $l = d = m \neq k$ or $l = k = m \neq d$, we have

$$\tilde{Q} = \bar{T} \mu^* S_6. \tag{63}$$

In the above expressions we define $\bar{P} = \mathbb{E}[X^2]$, $\bar{T} = \mathbb{E}[|X|^2 X]$. Hence, (54) reads

$$\begin{aligned}
\tilde{Q} &= |\mu|^4 S_2 + (2P^2 + |\bar{P}|^2) S_3 \\
&\quad + (4P|\mu|^2 + \bar{P}\mu^{*2} + \bar{P}^*\mu^2) S_4 \\
&\quad + Q S_5 + 2(\bar{T}\mu^* + \bar{T}^*\mu) S_6 \\
&= \frac{1}{3} \Bigg[Q + 4P(P - |\mu|^2) \\
&\quad + 2(|\bar{P}|^2 - \Re\{\bar{P}\mu^{*2}\}) + 2\Re\{\bar{T}\mu^*\} \Bigg].
\end{aligned} \tag{64, 65}$$

Expanding the terms $|\bar{P}|^2 - \Re\{\bar{P}\mu^{*2}\}$ and $\Re\{\bar{T}\mu^*\}$ in (65), we have

$$|\bar{P}|^2 - \Re\{\bar{P}\mu^{*2}\} = (P_r - P_i)(P_r - P_i - (\mu_r^2 - \mu_i^2)), \tag{66}$$

$$\Re\{\bar{T}\mu^*\} = \mu_r(T_r + \mu_r P_i) + \mu_i(T_i + \mu_i P_r). \tag{67}$$

Noting that $Q = Q_i + Q_r + 2P_r P_i$ and substituting in (65) along with (66) and (67), after some manipulations $\tilde{Q}$ reads

$$\tilde{Q} = \frac{1}{3}\big(Q_r + Q_i + 2(\mu_r T_r + \mu_i T_i)$$
$$+ 6(P_r P_i + P_r(P_r - \mu_r^2) + P_i(P_i - \mu_i))\big). \tag{68}$$

Substituting (68), (53) in (49) and substituting the result along with (42) in (37), and adding with (30) yields the result of the Proposition.

B. PROOF OF PROPOSITION 1

Note that constraining the input distributions $p_X(x)$ to those that are determined by their first and second moment statistics, the supremum in (15) is attained in general by a non-zero mean Gaussian distribution for each dimension, i.e., $\Re\{X\} \sim (\mu_r, \sigma_r^2)$ and $\Im\{X\} \sim (\mu_i, \sigma_i^2)$, where $\sigma_r^2 \triangleq P_r - \mu_r^2$ and $\sigma_i^2 \triangleq P_i - \mu_i^2$. Therefore, the optimization problem in (15) reads

$$\max_{\mu_r, \mu_i, P_r, P_i} \frac{f_w}{2}\big(\log(1 + a\sigma_r^2) + \log(1 + a\sigma_i^2)\big)$$

$$\text{s.t.} \quad \begin{cases} P_r + P_i \leq P_a \\ \alpha Q + \tilde{\alpha}\tilde{Q} + (\beta + \tilde{\beta})P + \gamma \geq P_d \\ \sigma_r^2 \geq 0, \ \sigma_i^2 \geq 0 \end{cases}, \tag{69}$$

where $a \triangleq 2|h|^2/f_w\sigma_w^2$. Writing the K.K.T. conditions for the optimization problem in (69), we have

$$\lambda_1(P_r + P_i - P_a) = 0, \ \lambda_1 \geq 0 \tag{70}$$

$$\lambda_2(\alpha Q + \tilde{\alpha}\tilde{Q} + (\beta + \tilde{\beta})P + \gamma - P_d) = 0, \ \lambda_2 \geq 0, \tag{71}$$

$$\zeta_r \sigma_r^2 = 0, \ \zeta_i \sigma_i^2 = 0, \ \zeta_r, \zeta_i \geq 0 \tag{72}$$

$$\zeta_r = \frac{-c_1 a}{1 + a\sigma_r^2} + \lambda_1$$
$$- \lambda_2(2(\alpha + \tilde{\alpha})(3P_r + P_i) + \beta + \tilde{\beta}), \tag{73}$$

$$\zeta_i = \frac{-c_1 a}{1 + a\sigma_i^2} + \lambda_1$$
$$- \lambda_2((\alpha + \tilde{\alpha})(3P_i + P_r) + \beta + \tilde{\beta}), \tag{74}$$

$$\frac{2c_1 a\mu_r}{1 + a\sigma_r^2} + 8\lambda_2(\alpha + \tilde{\alpha})\mu_r^3 + 2\zeta_r\mu_r = 0, \tag{75}$$

$$\frac{2c_1 a\mu_i}{1 + a\sigma_i^2} + 8\lambda_2(\alpha + \tilde{\alpha})\mu_i^3 + 2\zeta_i\mu_i = 0, \tag{76}$$

where $c_1 \triangleq (f_w \log e)/2$ and in (73) to (76) we used the following

$$\frac{\partial Q}{\partial P_r} = \frac{\partial \tilde{Q}}{\partial P_r} = 6P_l + 2P_i, \tag{77}$$

$$\frac{\partial Q}{\partial P_i} = \frac{\partial \tilde{Q}}{\partial P_i} = 6P_i + 2P_l, \tag{78}$$

$$\frac{\partial Q}{\partial \mu_r} = \frac{\partial \tilde{Q}}{\partial \mu_r} = -8\mu_r^3, \tag{79}$$

$$\frac{\partial Q}{\partial \mu_i} = \frac{\partial \tilde{Q}}{\partial \mu_i} = -8\mu_i^3. \tag{80}$$

It can be easily verified from (70), (73) and (74) that when $\lambda_2 = 0$, the maximum is achieved when $\mu_r = \mu_i = 0$ and $P_r = P_i = \frac{P_a}{2}$, yielding $P_{\text{dc,min}} = 2(\alpha + \tilde{\alpha})P_a^2 + (\beta + \tilde{\beta})P_a + \gamma$. For positive values of λ_2 from (73) it is verified that $\lambda_1 > 0$, which from (70) results that $P_r + P_i = P_a$. The condition $P_r + P_i = P_a$ reduces the number of variables P_i, P_r to one. Accordingly, since the mutual information $I(X;Y)$ is concave w.r.t. $P_i \in [0, P_a]$ attaining its maximum at $P_i = P_a/2$ and the delivered power P_{del} is convex w.r.t. $P_i \in [0, P_a]$ attaining its maximum at $P_i = 0$ or $P_i = P_a$ the Proposition is proved.

C. Proof of Proposition 3

To prove the series, we will use the following special cases of Riemann's zeta function and alternating series [15, Sec. 9.5]

$$\sum_{l=1}^{\infty} \frac{1}{l^2} = \frac{\pi^2}{6}, \tag{81}$$

$$\sum_{l=1}^{\infty} \frac{1}{l^4} = \frac{\pi^4}{90}, \tag{82}$$

$$\sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)^3} = \frac{\pi^3}{32}, \tag{83}$$

$$\sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)} = \frac{\pi}{4}. \tag{84}$$

We have

$$T_0 = \sum_l s_l \tag{85}$$

$$= \frac{1}{\pi} \sum_l \frac{(-1)^l}{\left(\frac{1}{2} + l\right)} \tag{86}$$

$$= \frac{2}{\pi} \left[\sum_{l=-\infty}^{-1} \frac{(-1)^l}{(2l+1)} + \sum_0^{\infty} \frac{(-1)^l}{(2l+1)} \right] \tag{87}$$

$$= \frac{2}{\pi} \left(\frac{\pi}{4} + \frac{\pi}{4} \right) = 1, \tag{88}$$

$$S_0 = \sum_l s_l^2 \tag{89}$$

$$= \sum_l \frac{(-1)^{2l}}{\pi^2 \left(\frac{1}{2} + l\right)^2} \tag{90}$$

$$= \frac{4}{\pi^2} \sum_l \frac{1}{(2l+1)^2} \tag{91}$$

$$= \frac{4}{\pi^2} \left[\sum_{l=-\infty}^{-1} \frac{1}{(2l+1)^2} + \sum_0^\infty \frac{1}{(2l+1)^2} \right] \tag{92}$$

$$= \frac{4}{\pi^2} \left(\frac{\pi^2}{8} + \frac{\pi^2}{8} \right) = 1, \tag{93}$$

$$T_1 = \sum_l s_l^3 \tag{94}$$

$$= \sum_l \frac{(-1)^{3l}}{\pi^3 \left(\frac{1}{2} + l\right)^3} \tag{95}$$

$$= \frac{8}{\pi^3} \left[\sum_{l=-\infty}^{-1} \frac{1}{(2l+1)^3} + \sum_0^\infty \frac{1}{(2l+1)^3} \right] \tag{96}$$

$$= \frac{8}{\pi^3} \left(\frac{\pi^3}{32} + \frac{\pi^3}{32} \right) = \frac{1}{2}, \tag{97}$$

$$S_5 = \sum_l s_l^4 \tag{98}$$

$$= \sum_l \frac{(-1)^{4l}}{\pi^4 \left(\frac{1}{2} + l\right)^4} \tag{99}$$

$$= \frac{16}{\pi^4} \sum_l \frac{1}{(2l+1)^4} \tag{100}$$

$$= \frac{16}{\pi^4} \left[\sum_{l=-\infty}^{-1} \frac{1}{(2l+1)^4} + \sum_0^\infty \frac{1}{(2l+1)^4} \right] \tag{101}$$

$$= \frac{16}{\pi^4} \left(\frac{\pi^4}{96} + \frac{\pi^4}{96} \right) = \frac{1}{3}, \tag{102}$$

$$S_1 = \sum_l \sum_{k,k \neq l} s_l s_k \tag{103}$$

$$= \sum_l s_l \left(\sum_k s_k - s_l \right) \tag{104}$$

$$= \left(\sum_l s_l \right)^2 - \sum_l s_l^2 \tag{105}$$

$$= 1 - 1 = 0, \tag{106}$$

$$S_3 = \sum_l \sum_{k,k \neq l} s_l^2 s_k^2 \tag{107}$$

$$= \sum_l s_l^2 \left(\sum_k s_k^2 - s_l^2 \right) \tag{108}$$

$$= \left(\sum_l s_l^2 \right)^2 - \sum_l s_l^4 \tag{109}$$

$$= 1 - \frac{1}{3} = \frac{2}{3}, \tag{110}$$

$$S_6 = \sum_l \sum_{k,k \neq l} s_l^3 s_k \tag{111}$$

$$= \sum_l s_l^3 \left(\sum_k s_k - s_l \right) \tag{112}$$

$$= \frac{1}{2} - \frac{1}{3} = \frac{1}{6}, \tag{113}$$

$$S_4 = \sum_l \sum_{k,k \neq l} \sum_{\substack{d,d \neq l \\ d \neq k}} s_l^2 s_k s_d \tag{114}$$

$$= \sum_l \sum_{k,k \neq l} s_l^2 s_k \left(\sum_d s_d - s_l - s_k \right) \tag{115}$$

$$= \sum_l s_l^2 \left((1 - s_l) \sum_{k,k \neq l} s_k - \sum_{k,k \neq l} s_k^2 \right) \tag{116}$$

$$= \sum_l s_l^2 \left((1 - s_l)^2 - (1 - s_l^2) \right) \tag{117}$$

$$= \sum_l 2 s_l^2 (s_l^2 - s_l) \tag{118}$$

$$= 2 \left(\frac{1}{3} - \frac{1}{2} \right) = -\frac{1}{3}, \tag{119}$$

$$S_2 = \sum_l \sum_{k,k \neq l} \sum_{\substack{d,d \neq l \\ d \neq k}} \sum_{\substack{m,m \neq d \\ m \neq l \\ m \neq k}} s_l s_k s_d s_m \tag{120}$$

$$= \sum_l \sum_{k,k \neq l} \sum_{\substack{d,d \neq l \\ d \neq k}} s_l s_k s_d (1 - s_d - s_l - s_k) \tag{121}$$

$$= \sum_l \sum_{k,k \neq l} s_l s_k \left((1 - s_l - s_k) \sum_{\substack{d,d \neq l \\ d \neq k}} s_d - \sum_{\substack{d,d \neq l \\ d \neq k}} s_d^2 \right) \tag{122}$$

$$= \sum_l \sum_{k,k \neq l} s_l s_k \left((1 - s_l - s_k)^2 - (1 - s_l^2 - s_k^2) \right) \tag{123}$$

$$= \sum_l s_l \left(2s_l(s_l - 1)(1 - s_l) + \sum_{k, k \neq l} 2s_k(s_k^2 + s_l s_k - s_k) \right) \tag{124}$$

$$= \sum_l s_l(-6s_l^3 + 6s_l^2 - 1)) \tag{125}$$

$$= -\frac{6}{3} + \frac{6}{3} - 1 = 0. \tag{126}$$

REFERENCES

[1] L. R. Varshney, "Transporting information and energy simultaneously," in *IEEE Int. Sym. Inf. Theory*, Jul. 2008, pp. 1612–1616.

[2] P. Grover and A. Sahai, "Shannon meets tesla: Wireless information and power transfer," in *IEEE Int. Sym. Inf. Theory*, Jun. 2010, pp. 2363–2367.

[3] R. Zhang and C. Keong, "Mimo broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[4] 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.

[5] ——, "Joint wireless information and energy transfer in a k-user mimo interference channel," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5781–5796, Oct. 2014.

[6] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3622–3636, Jul. 2013.

[7] Y. Huang and B. Clerckx, "Relaying strategies for wireless-powered mimo relay networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6033–6047, Sep. 2016.

[8] 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.

[9] B. Clerckx and E. Bayguzina, "Waveform design for wireless power transfer," *IEEE Trans. Signal Processing*, vol. 64, no. 23, pp. 6313–6328, Dec. 2016.

[10] B. Clerckx, "Waveform and transceiver design for simultaneous wireless information and power transfer," *CoRR*, vol. abs/1607.05602, 2016. [Online]. Available: http://arxiv.org/abs/1607.05602

[11] ——, "Waveform optimization for swipt with nonlinear energy harvester modeling," in *20th International ITG Workshop Smart Antennas (WSA)*, Mar. 2016, pp. 1–5.

[12] R. P. Feynman, *Feynman Lectures on Computation*, J. G. Hey and R. W. Allen, Eds.   Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1998.

[13] R. Landauer, *Computation, Measurement, Communication and Energy Dissipation*.   IBM Thomas J. Watson Research Division, 1987.

[14] M. P. Frank, *Reversibility for efficient computing*.   Ph.D. dissertation, Massachusetts Institute of Technology, 1999.

[15] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*.   Academic Press, 1965.