SWIPT Signalling over Frequency-Selective Channels with a Nonlinear Energy Harvester: Non-Zero Mean and Asymmetric Inputs

Morteza Varasteh[†], Borzoo Rassouli[*] and Bruno Clerckx[†]

[†] Department of Electrical and Electronic Engineering, Imperial College London, UK.

[*] School of Computer Science and Electronic Engineering, University of Essex, UK.

{m.varasteh12; b.clerckx}@imperial.ac.uk, b.rassouli@essex.ac.uk.

Abstract—Simultaneous Wireless Information and Power Transmission (SWIPT) over a point-to-point frequency-selective channel is studied. Leveraging a small-signal model for a nonlinear energy harvester, a general form of delivered power in terms of system baseband parameters is derived, which demonstrates the dependency of the delivered power on higher moments of the baseband channel input. The optimization problem of maximizing Rate-Power (RP) region is studied. Assuming that the Channel State Information (CSI) is available at both the receiver and the transmitter, and constraining to non-zero mean Gaussian input distributions, an optimization algorithm for power allocation among different subchannels is studied. As a special case, optimality conditions for zero mean Gaussian inputs are derived. Results obtained from numerical optimization demonstrate the superiority of non-zero mean Gaussian inputs (with asymmetric power allocation in each complex subchannel) in yielding a larger RP region compared to their zero mean and non-zero mean (with symmetric power allocation in each complex subchannel) counterparts. The results motivate the design of new modulation for SWIPT and contrast with SWIPT design under linear energy harvesting, for which circularly symmetric Gaussian inputs with water-filling power allocation are optimal.

I. INTRODUCTION

Radio-Frequency (RF) waves can be utilized for transmission of both information and power simultaneously. RF transmissions of these quantities have traditionally been treated separately. Currently, the community is experiencing a paradigm shift in wireless network design, namely unifying transmission of information and power, so as to make the best use of the RF spectrum and radiation, as well as the network infrastructure for the dual purpose of communicating and energizing [1]. This has led to a growing interest in integrating energy harvesting technologies into communication networks, resulting in the emerging area of Simultaneous Wireless Information and Power Transmission (SWIPT). As one of the primary works in the information theory literature, Varshney studied SWIPT in [2], in which he characterized the capacity-power function for

a point-to-point discrete memoryless channel. 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 [3], MIMO broadcasting [4], interference channel [5].

The ambient energy sources make energy harvesting unreliable for applications where a constant flow of energy is of significant importance. An energy harvesting technology that overcomes this limitations is Wireless Power Transfer (WPT), where dedicated signals are designed to transmit the required power at the user end. The main challenge in WPT is to increase the Direct-Current (DC) power at the output of the harvester, known as rectenna, without increasing transmit power. A rectenna is made of a nonlinear device followed by a lowpass filter to extract a DC power out of an RF input signal. The amount of DC power collected is a function of the input power level and the RF-to-DC conversion efficiency. The RF-to-DC conversion efficiency is not only a function of the rectenna design but also of its input waveform (power and shape) [6]–[10]. Consequently, the conversion efficiency is not a constant but a nonlinear function of the input waveform (power and shape). In [6], [10], it is shown that the RF-to-DC conversion efficiency is a function of the rectenna's structure, as well as its input waveform (power and shape). Accordingly, in order to maximize the rectenna's output power, a systematic waveform design is crucial to make the best use of an available RF spectrum [10]. In [10], an analytical model for the rectenna's output is introduced via the Taylor expansion of the diode characteristic function and a systematic design for multisine waveform is derived. The nonlinear model and the design of the waveform was validated using circuit simulations in [10], [11] and recently confirmed through prototyping and experimentation in [12], [13]. Those works also confirm the inaccuracy of the linear dependence of the rectifier's output power on its input signal (power and shape)[1]. As one of the main conclusions, it is shown that the rectifier's nonlinearity is beneficial to the system performance and has a significant

---

[1]The linear model has for consequence that the RF-to-DC conversion efficiency of the energy harvester (EH) is constant and independent of the harvester's input waveform (power and shape) [4], [14].

impact on the design of signals and systems involving wireless power transmission.

SWIPT can potentially bring significant gains in information and power transmission, such as spectral efficiency and energy consumption. For example in Tactile Internet, Cyber Physical Systems (CPS) and Internet of Things (IoT), SWIPT technologies can be of importance when it comes to supply energy remotely in addition to the exchange of information among different nodes. The SWIPT literature has so far, to a great extent, ignored the nonlinearity of the EH and has focused on the linear model of the rectifier, e.g., [3]–[5]. However, it is recognized that considering the harvester nonlinearity changes the design of SWIPT at the physical layer and medium access control layer [1]. Nonlinearity leads to various energy harvester models [9], [10], [15], new designs of modulation and input distribution [16]–[18], waveform [8], RF spectrum use [8], transmitter and receiver architecture [8], [17], [19] and resource allocation [9], [20], [21]. Of particular interest is the role played by nonlinearity on SWIPT signalling in single-carrier and multi-carrier transmissions [1], [8], [16], [17], [22]. In multi-carrier transmissions, it is shown in [8] that inputs modulated according to the Circular Symmetric Complex Gaussian (CSCG) distributions, improve the delivered power compared to an unmodulated continuous waves. Furthermore, in [16], it is shown that for an Additive White Gaussian Noise (AWGN) channel with complex Gaussian inputs under average power and delivered power constraints, depending on the receiver demand of information and power, the power allocation between real and imaginary components is asymmetric. As an extreme point, when the receiver merely demands for power, all the transmitter power budget is allocated to either real or imaginary components. In [17], [22], it is shown that the capacity achieving input distribution of an AWGN channel under average, peak and delivered power constraints is discrete in amplitude with a finite number of mass-points and with a uniformly distributed independent phase. In multi-carrier transmission, however, it is shown in [8] that non-zero mean Gaussian input distributions lead to an enlarged Rate-Power (RP) region compared to CSCG input distributions[2]. This highlights that the choice of a suitable input (and therefore modulation and waveform) for SWIPT is affected by the EH nonlinearity and motivates the study of AWGN channels under nonlinear power constraints.

Our interests in this paper lie in the apparent difference in input distribution for single-carrier and multi-carrier transmission, that is single-carrier favors asymmetric inputs [16], while multi-carrier favors non-zero mean inputs [8]. We aim at tackling the design of channel input for SWIPT under nonlinear constraints using a unified framework based on non-zero mean and asymmetric distributions. To that end, we study SWIPT in a multi-carrier setting subject to nonlinearities of the EH. We consider a frequency-selective channel subject to transmit average power and receiver delivered power constraints. We

mainly focus on complex Gaussian inputs, where inputs of each real subchannel are independent of each other and on each real subchannel the inputs are independent and identically distributed (i.i.d.).

We aim at reconciling the two main observations of the previous paragraph: that is, outperforming of asymmetric Gaussian inputs and non-zero mean Gaussian inputs compared to CSCG inputs in single-carrier transmission [16] and multi-carrier transmission [8], respectively. The contributions of this paper are listed below.

- First, taking the advantage of the small-signal approximation for rectenna introduced in [8], we obtain the general form of the delivered power in terms of system baseband parameters. It is shown that, first, the delivered power at the receiver is dependent on higher moments of the channel input, such as the first, second and forth moments. This contrasts with SWIPT signal design under linear EH assumption, where CSCG Gaussian inputs are optimal [3]. Second, the amount of delivered power on each subchannel is dependent on its adjacent subchannels.

- Assuming non-zero mean Gaussian inputs, an optimization algorithm is introduced. Relying on numerical optimizations two observations are made for the scenarios where the receiver is interested in both information and power, simultaneously: first, allowing the input to be non-zero mean improves the rate-power region, significantly, and second, for receiver demands concerning information and power, the power allocation between real and imaginary components of each complex subchannel is asymmetric in general. These results can be thought of as generalization of the results in [8], [16], where asymmetric power allocation (in flat fading channels) and non-zero mean inputs (in frequency-selective channels) are proposed, respectively, in order to achieve larger RP regions. The results also reconcile the seemingly different strategies of [8], [17] and show that a joint optimization of the mean and the asymmetric power allocation for Gaussian inputs in a general frequency-selective channel can enlarge the rate-power region. This is one more step toward identifying the optimal input distributions for SWIPT in frequency selective channels subject to EH nonlinearity[3].

- As a special scenario, we consider the optimized zero mean Gaussian inputs under the assumption of nonlinear EH. For this scenario, optimality conditions are derived. It is shown that (similar to non-zero mean inputs) under nonlinear assumption for the EH, the power allocation on each subchannel is dependent on other subchannels as well. Forcing the optimality conditions to be satisfied (numerically), it is observed that a larger RP region is obtained in contrast to the optimal zero mean inputs under the linear assumption for the EH.

An important implication of the results here is that replacing the conventional modulations, such as BPSK, QPSK, 16QAM,

---

[2]The non-zero mean channel input has also revealed benefits in other SWIPT structures such as [23] for two user Gaussian interference channel, though for different reasons not motivated by the EH nonlinearity.

[3]The optimal input distribution in the frequency flat case has been characterized in [17]

with (in general) non-zero mean and in an asymmetric fashion (that is the power allocation across real and imaginary components of the modulations in each subchannel are not equal) can lead to higher delivered power at the receiver.

Organization: In Section II, we introduce the system model. In Section III, the studied problem is introduced. In Section IV, the delivered power at the output of the EH is obtained in terms of system baseband parameters. In Section V, the rate-power maximization over frequency-selective channels with non-zero mean Gaussian inputs is considered. As a special case, the optimality conditions for power allocation on different subchannels are obtained for zero mean Gaussian inputs. In Section VI, WPT and SWIPT optimization for the studied problem is introduced and numerical results are presented. We conclude the paper in Section VII and the proofs for some of the results are provided in the Appendices at the end of the paper.

Notation: Throughout this paper, random variables and their realizations are represented by bold-small and small letters, respectively. $\mathbb{E}[Y(t)]$ and $\mathcal{E}[Y(t)]$ denote the expectation over statistical randomness and the average over time of the process $Y(t)$, respectively, i.e.,

$$\mathbb{E}[Y(t)] = \int_{-\infty}^{\infty} y(t)\phi_{Y(t)}dy, \tag{1}$$

$$\mathcal{E}[Y(t)] = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} Y(t)dt, \tag{2}$$

where $\phi_{Y(t)}(y)$ denotes the Probability Distribution Function (PDF) of the process $Y(t)$. $\otimes$ denotes circular convolution. The standard CSCG distribution is denoted by $\mathcal{CN}(0,1)$. Complex conjugate of a complex number c is denoted by c^*. $\mathfrak{R}\{\cdot\}$ and $\mathfrak{I}\{\cdot\}$ are real and imaginary operators, respectively. For a complex random variable $\mathbf{v}$, we define $Q \triangleq \mathbb{E}[|\mathbf{v}|^4]$, $P \triangleq \mathbb{E}[|\mathbf{v}|^2]$, $\bar{P} \triangleq \mathbb{E}[\mathbf{v}^2]$, $\mu \triangleq \mathbb{E}[\mathbf{v}]$ and $\sigma^2 \triangleq \mathbb{E}[|\mathbf{v} - \mu|^2]$. The moments corresponding to real and imaginary components of $\mathbf{v}$ are represented by subscripts r and i, respectively, i.e., $Q_r \triangleq \mathbb{E}[\mathfrak{R}\{\mathbf{v}\}^4]$, $P_r \triangleq \mathbb{E}[\mathfrak{R}\{\mathbf{v}\}^2]$, $\mu_r \triangleq \mathbb{E}[\mathfrak{R}\{\mathbf{v}\}]$ and $\sigma_r^2 \triangleq \mathbb{E}[|\mathfrak{R}\{\mathbf{v}\} - \mu_r|^2]$ and similarly for imaginary counterparts. $(\cdot)_N$ denotes remainder of the argument with respect to N. $\delta_k = 1$ for $k = 0$ and zero elsewhere. $\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$ and $\delta_k^l \triangleq 1 - \delta_{l-k}$. f^x denotes the partial derivative of the function f with respect to x, i.e., $\frac{\partial f}{\partial x}$. The vector $[\mathbf{v}_0, \ldots, \mathbf{v}_{N-1}]$ is represented by $\mathbf{v}^N$. Throughout the paper, complex subchannels and their real/imaginary components are referred to as c-subchannels and r-subchannels, respectively. The subscript l denotes the lth subchannel. For example, P_l and P_{lr} denote the average power of the lth c-subchannel and the average power of the real component of the lth c-subchannel, respectively. When summation over an index is considered, we neglect the lower and upper limits of the summation if they are $-\infty$ and ∞, respectively, i.e., $\sum_n \triangleq \sum_{n=-\infty}^{\infty}$.

II. SYSTEM MODEL

Considering a point-to-point L-tap frequency-selective channel, in the following, we explain the operation of the transmitter and the receiver.

A. Transmitter

The transmitter utilizes Orthogonal Frequency Division Multiplexing (OFDM) to transmit information and power over the channel. Let $\mathbf{v}^N$ denote the modulated Information-Power (IP) complex symbols over N sub-carriers (c-subchannels), occupying the overall bandwidth of f_w Hz and being uniformly separated by f_w/N Hz. Inverse Discrete Fourier Transform (IDFT) is applied over IP symbols $\mathbf{v}^N$ and Cyclic Prefix (CP) is added to produce the time domain signal $X[n]$ given by

$$X[n+L] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \mathbf{v}_k e^{\frac{j2\pi nk}{N}}, \; n = 0, \ldots, N-1. \tag{3}$$

Next, the signal

$$X(t) = \sum_{n=0}^{N+L-1} X[n]\mathrm{sinc}(f_w t - n), \tag{4}$$

is upconverted to the carrier frequency f_c and is transmitted over the channel.

B. Receiver

The filtered received RF waveform at the receiver is modelled as

$$Y_{\text{rf}}(t) = \sqrt{2}\mathfrak{R}\left\{Y(t)e^{j2\pi f_c t}\right\}, \tag{5}$$

where $Y(t)$ is the baseband equivalent of the channel output with bandwidth $[-f_w/2, f_w/2]$ Hz. In order to guarantee narrowband transmission, we assume that $f_c \gg 2f_w$.

Delivered Power: The power of the signal $Y_{\text{rf}}(t)$ (denoted by P_{dc}) is harvested using a rectenna. The delivered power is modelled as

$$P_{\text{dc}} = \mathbb{E}[\mathcal{E}[k_2 Y_{\text{rf}}(t)^2 + k_4 Y_{\text{rf}}(t)^4]], \tag{6}$$

where k_2 and k_4 are constants. The model in (6) is based on truncating (up to the 4th moment) of the diode characteristic function in the rectenna. Note that due to zero time average of odd moments of $Y_{\text{rf}}(t)$, i.e., $\mathcal{E}[Y_{\text{rf}}(t)] = 0$, we only have the even moments in (6). The reader is referred to [10] for detailed explanations of the model. Also note that according to [8], rectenna's output is in the form of current with unit Ampere. However, since the common terminology in the SWIPT literature is power, such as "information-power tradeoff" and since power is proportional to current, with some abuse of notation, we refer to the term in (6) as power.

Information Receiver: The process $Y_{\text{rf}}(t)$ is downconverted and sampled with sampling frequency f_w producing $Y[m] \triangleq Y(m/f_w)$ given by

$$Y[m] = \sum_{d=0}^{L-1} \tilde{h}_d X[m-d] + Z[m], \; m = L, \ldots, N+L-1, \tag{7}$$

where $Z[m]$ represents the additive noise process at time $t = m/f_w$. $\tilde{h}_d$ is the dth c-subchannel tap and $X[m-d]$ is the process $X(t)$ given in (4) at time $(m-d)/f_w$.

Considering one OFDM block, the receiver discards the CP and converts the N symbols back to the frequency domain by applying DFT on (7), such that

$$\mathbf{y}_l = h_l \mathbf{v}_l + \mathbf{w}_l, \quad l = 0, \cdots, N-1, \tag{8}$$

where $\mathbf{y}_l, l = 0, \cdots, N-1$ is the DFT of $Y[m], m = L, ..., L + N - 1$. h_l, $\mathbf{v}_l$ and $\mathbf{w}_l$ are DFTs of the extended channel vector $\tilde{H} \triangleq [\tilde{h}_0, \cdots, \tilde{h}_{L-1}, 0, \cdots, 0]_{1 \times N}$, symbols $X[m], m = L, ..., L + N - 1$ (equivalently, samples of $X(t)$ at times m/f_w) and noise samples $Z[m], m = L, ..., L + N - 1$, respectively. That is,

$$h_l = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \tilde{H}[n] e^{-\frac{j2\pi nl}{N}}, \quad l = 0, \cdots, N-1, \tag{9a}$$

$$\mathbf{v}_l = \frac{1}{\sqrt{N}} \sum_{n=L}^{L+N-1} X[n] e^{-\frac{j2\pi nl}{N}}, \quad l = 0, \cdots, N-1, \tag{9b}$$

and similarly for $\mathbf{w}_l$, $l = 0, \cdots, N-1$. We assume $\mathbf{w}_l$, $l = 0, \cdots, N-1$ as i.i.d. and CSCG random variables with variance σ_w^2, i.e., $\mathbf{w}_l \sim \mathcal{CN}(0, \sigma_w^2)$ for $l = 0, \cdots, N-1$. The channel frequency response is assumed to be known at the transmitter and the receiver.

III. PROBLEM STATEMENT

We aim at maximizing the rate of transmitted information under an average power constraint at the transmitter and a delivered power constraint at the receiver, given that the input in each c-subchannel $l = 0, \dots, N-1$ is distributed according to a non-zero mean complex Gaussian distribution. We also assume that in each c-subchannel the real and imaginary components are independent. Accordingly, the optimization problem consists in the maximization of the mutual information between the channel input $\mathbf{v}^N$ and the channel output $\mathbf{y}^N$ (see eq. 8) under an average power constraint at the transmitter and a delivered power constraint at the receiver, such that, $\mathbf{v}_{lr} \sim \mathcal{N}(\mu_{lr}, P_{lr} - \mu_{lr}^2)$ and $\mathbf{v}_{li} \sim \mathcal{N}(\mu_{li}, P_{li} - \mu_{li}^2)$ for $l = 0, \cdots, N-1$. Hence, we have

$$\max_{\mu_{lr}, \mu_{li}, P_{lr}, P_{li}, \ l=0,...N-1} I\left(\mathbf{v}^N; \mathbf{y}^N\right)$$
$$\text{s.t.} \quad \begin{cases} \sum_{l=0}^{N-1} P_l \leq P_a \\ P_{\text{dc}} \geq P_d \end{cases}, \tag{10}$$

where $P_l = P_{lr} + P_{li}$ and $\mu_l = \mu_{lr} + j\mu_{li}$. P_a is the available power budget at the transmitter. P_d is the minimum amount of average delivered power at the receiver. Maximization is taken over all the means μ_{lr}, μ_{li} and powers P_{lr}, P_{li} ($l = 0, \dots, N-1$) of independent complex Gaussian inputs $\mathbf{v}^N$, such that the constraints in (10) are satisfied.

IV. POWER METRIC IN TERMS OF CHANNEL BASEBAND PARAMETERS

In this section, we study the delivered power at the receiver based on the model in (6). Note that most of the communication processes, such as, coding/decoding, modulation/demodulation, etc, are done at the baseband. Therefore, from a communication system design point of view, it is most preferable to have baseband equivalent representation of the system. Henceforth, in the following Proposition, we derive the delivered power in (6) at the receiver in terms of system baseband parameters. For brevity of representation, we neglect the delivered power from CP in our calculations, and also we assume that N is odd (calculations can be easily extended to even values of N, following similar steps). The following proposition, expresses the delivered power P_{dc} in (6) in terms of the channel and its input baseband parameters.

Proposition 1. *Given that the inputs on each r-subchannel are i.i.d. and that the inputs on different r-subchannels are independent, the delivered power P_{dc} at the receiver can be expressed in terms of the channel baseband parameters and statistics of the channel input distribution as*

$$P_{dc} = \sum_{l=0}^{N-1} \left\{ \alpha_l Q_l + \left(\beta_l + g(P_l)\right) P_l + \eta \right.$$
$$+ \Re \left\{ \bar{P}_l \sum_{k=1}^{\frac{N-1}{2}} \mu_{(l+k)_N}^* \mu_{(l-k)_N}^* \Phi_{l,k} \right\}$$
$$+ \delta_{(N-1)}^l \cdot \sum_{k=1}^{\frac{N-1}{2}} \sum_{\substack{m=l+1 \\ m \neq (l+k)_N \\ m \neq (l-k)_N}}^{N-1} \Re \left\{ \mu_l \mu_m \mu_{(l-k)_N}^* \mu_{(m-k)_N}^* \Psi_{l,m,k} \right\} \right\}$$

$$\triangleq \sum_{l=0}^{N-1} f_{ib}(Q_l, P_l, \bar{P}_l, \mu_l, h_l, N), \tag{11}$$

where N is odd and α_l, β_l, $\gamma_{l,m}$, η, $\Phi_{l,k}$, $\Psi_{l,n,k}$ and $g(P_l)$ are defined as

$$\alpha_l = \frac{3k_4}{4N}(|h_l|^4 + |h_l^u|^4), \tag{12a}$$

$$\beta_l = k_2|h_l|^2 + 3k_4\sigma_w^2 \left(|h_l|^2 + |h_l^u|^2\right), \tag{12b}$$

$$\gamma_{m,l} = \frac{3k_4}{N}(|h_l|^2|h_m|^2 + |h_l^u|^2|h_m^u|^2), \tag{12c}$$

$$\eta = k_2\sigma_w^2 + 3Nk_4\sigma_w^4, \tag{12d}$$

$$\Phi_{l,k} = \frac{3k_4}{2N}\left(h_l^2 h_{(l+k)_N}^* h_{(l-k)_N}^* + h_l^{u2} h_{(l+k)_N}^{u*} h_{(l-k)_N}^{u*}\right), \tag{12e}$$

$$\Psi_{l,m,k} = \frac{3k_4}{N}\left(h_l h_m^* h_{(l-k)_N}^* h_{(m-k)_N} + h_l^u h_m^{u*} h_{(l-k)_N}^{u*} h_{(m-k)_N}^u\right), \tag{12f}$$

$$g(P_l) = \delta_{N-1}^l \sum_{m=l+1}^{N-1} \gamma_{m,l} P_m, \tag{12g}$$

with h_l^u, $l = 0, \cdots, N-1$ being the samples of the channel at times between two consecutive information samples (for more details see Appendix A-B).

Proof: See Appendix A.

Remark 1. Note that (as also mentioned in Proposition 1) the delivered power is based on the following assumptions: first, the inputs on different r-subchannels are independent. Second, the input on each r-subchannel is i.i.d.. Obtaining a closed form expression for the delivered power P_{dc} at the receiver when the inputs on different r-subchannels 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 among the inputs of different r-subchannels. As another point, we note that in the calculations for the delivered power in Proposition 1, we neglect the delivered power from CP for brevity. This along with the aforementioned assumptions on the input distributions, bears the fact that the real delivered power (based on the introduced model in (6)) is larger than (11). Indeed, the subscript ib in (11) stands for inner bound in order to express this point.

Remark 2. Note that similar results in [16] are reported for single-carrier AWGN channel, where the delivered power is dependent on higher moments of the channel input. In [8], superposition of deterministic and CSCG signals are assumed for multi-carrier transmissions with the assumption that the receiver utilizes power splitter. Part of the signal is used for power transmission and the other part is used for information transmissions. In comparison to the results in [8], we note that, first, a different transmission model is considered here. Second, the channel input is generalized in the sense that it allows asymmetric power allocation across all r-subchannels. Also, (due to simplifying the calculations) at the receiver, no power splitter is assumed. However, the results can be simply generalized to a scenario where a power splitter is present at the receiver.

V. RATE-POWER MAXIMIZATION OVER GAUSSIAN INPUTS

In this section, we consider the SWIPT optimization problem in (10). We obtain the optimality conditions in their general form (assuming non-zero mean inputs) to be used in Section VI in order to obtain (locally) optimal power allocations for different r-subchannels. In order to better understand the problem, the optimality conditions are specialized for zero mean Gaussian inputs.

A. SWIPT with non-zero mean complex Gaussian inputs

Assuming that the inputs of c-subchannels v^N are in general with non-zero mean, the problem in (10) can be rewritten as follows

$$\max_{\substack{P_{lr}, P_{li}, \mu_{lr}, \mu_{li} \\ l=0,...,N-1}} \sum_{l=0}^{N-1} c_0\big(\log(1 + a_l\sigma_{lr}^2) + \log(1 + a_l\sigma_{li}^2)\big)$$

$$\text{s.t.} \begin{cases} \sum_{l=0}^{N-1} P_l \leq P_a \\ \sum_{l=0}^{N-1} f_{ib}(P_l, \bar{P}_l, \mu_l, h_l, N) \geq P_d \\ \sigma_{lr}^2 \geq 0, \sigma_{li}^2 \geq 0, \ l = 0,...,N-1 \end{cases}, \quad (13)$$

where $c_0 = \frac{f_w}{2N}$, $a_l = \frac{2N|h_l|^2}{f_w \sigma_w^2}$, $\sigma_{lr}^2 = P_{lr} - \mu_{lr}^2$, $\sigma_{li}^2 = P_{li} - \mu_{li}^2$. Note that for a Gaussian distribution in the function $f_{ib}(\cdot)$, we have $Q_l = 3(P_{lr}^2 + P_{li}^2) - 2(\mu_{li}^4 + \mu_{lr}^4) + 2P_{lr}P_{li}$, $\bar{P}_l = P_{lr} - P_{li} + 2j\mu_{lr}\mu_{li}$.

In Section VI, we consider the numerical optimization of problem (13) by considering its Lagrangian. Unfortunately, it is cumbersome to derive analytical results on the optimal solution of problem (13). However, by obtaining the KKT conditions, it can be shown that for the optimal solution, the average power

constraint is satisfied with equality (see Appendix B for the details).

As explained in Section VI, numerical results reveal that non-zero mean asymmetric complex Gaussian inputs result in larger RP region compared to their zero mean counterparts. However, in order to better understand the problem in its general form (assuming non-zero mean), it is beneficial to look into the optimality conditions of zero mean inputs.

B. SWIPT with zero mean complex Gaussian inputs

In the following, we obtain the optimality conditions for power allocation among different r-subchannels, when the input distributions are complex Gaussian with zero mean and with independent components.

Lemma 1. If $\mu^N = 0^N$, the optimal power allocation $P_r^{N\star}$, $P_i^{N\star}$ for problem (13) satisfies the average power and delivered power constraints with equality, i.e.,

$$\sum_{l=0}^{N-1} P_l^{\star} = P_a, \quad (14a)$$

$$\sum_{l=0}^{N-1} f_{ib}(P_l^{\star}, \bar{P}_l^{\star}, 0, h_l, N) = P_d, \quad (14b)$$

with $f_{ib}(P_l^{\star}, \bar{P}_l^{\star}, 0, h_l, N) = \alpha_l Q_l^{\star} + \big(\beta_l + g(P_l^{\star})\big)P_l^{\star} + \eta$. Also for the optimal vectors $P_r^{N\star}$, $P_i^{N\star}$ we have

$$P_{lr}^{\star} \cdot \big(\lambda_1 - G_l(P_r^{N\star}, P_i^{N\star})\big) = 0, \ l = 0,...,N-1, \quad (15a)$$

$$P_{li}^{\star} \cdot \big(\lambda_1 - G_l(P_i^{N\star}, P_r^{N\star})\big) = 0, \ l = 0,...,N-1, \quad (15b)$$

with

$$G_l(P_r^N, P_i^N) \triangleq \frac{c_1 a_l}{1 + a_l P_{lr}} + 6\lambda_2 \alpha_l P_{lr} + \lambda_2(2\alpha_l P_{li} + \beta_l + g_1(P_l)), \quad (16)$$

for some

$$\lambda_1 \geq \max_{l=0,...,N-1}\{G_l(P_r^{N\star}, P_i^{N\star}), G_l(P_i^{N\star}, P_r^{N\star})\}, \quad (17a)$$

$$\lambda_2 \geq 0, \quad (17b)$$

and $g_1(P_l) \triangleq \sum_{\substack{m=0 \\ m \neq l}}^{N-1} \gamma_{m,l} P_m$. For $\lambda_2 = 0$, the optimal power allocations are simplified to water-filling solution, i.e.,

$$P_{lr}^{\star} = P_{li}^{\star} = \max\left\{0, \frac{c_1}{\lambda_1} - \frac{1}{a_l}\right\}, \ for \ l = 0,\cdots,N-1. \quad (18)$$

Proof: See Appendix C.

Remark 3. Note that the delivered power in the lth c-subchannel for zero mean Gaussian inputs, i.e.,

$$f_{ib}(P_l^{\star}, \bar{P}_l^{\star}, 0, h_l, N) = \alpha_l Q_l^{\star} + \big(\beta_l + g(P_l^{\star})\big)P_l^{\star} + \eta,$$

is dependent on other c-subchannels through $g(P_l^{\star})$. This is in contrast with the linear model of the EH, i.e., $P_{dc} \propto \mathbb{E}\mathcal{E}[Y_{rf}(t)^2]$, where the delivered power is obtained as $|h_l|^2 P_l + \sigma_w^2$. Also, it can be easily verified that for zero mean inputs with nonlinear
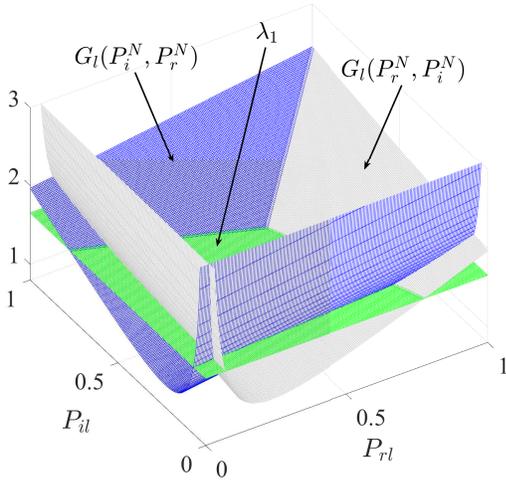
Figure 1: Illustration of intersection λ_1 with the functions $G_l(P_r^N, P_i^N)$ and $G_l(P_i^N, P_r^N)$, defined in (16) under average power constraint $P_a = 1$.

EH, $P_{lr} = P_{lr}^\star, P_{li} = P_{li}^\star$ yields the same delivered power/ transmitted information as $P_{lr} = P_{li}^\star, P_{li} = P_{lr}^\star$.

The optimality conditions of Lemma 1 in (15) can be interpreted as follows. The functions $G_l(P_r^N, P_i^N)$, $l = 0, \ldots, N-1$ are positive and convex (the Hessian matrix is positive definite). Also note that $G_l(P_i^N, P_r^N)$ is a mirrored version of $G_l(P_r^N, P_i^N)$ with respect to the surface $P_{lr} = P_{li}$. Assume that $\lambda_2 > 0$ is given and that λ_1 is chosen as a large value (so that it satisfies (17a)). Consider the intersection of the horizontal surface λ_1 with functions $G_l(P_r^N, P_i^N)$ and $G_l(P_i^N, P_r^N)$ for some index l. Depending on the value of λ_1 and shape of the functions G_l, different pairs of (P_{lr}, P_{li}) satisfy simultaneously

$$\lambda_1 = G_l(P_r^N, P_i^N) = G_l(P_i^N, P_r^N). \quad (19)$$

The number of these solution pairs (P_{lr}, P_{li}) for each index l can be verified to vary from three to four. That is, if $\lambda_1 > G_l(0^N, 0^N)$, there are three solutions, and if $\lambda_1 \leq G_l(0^N, 0^N)$, there are four solutions for (19). In Figure 1, an illustration of the intersection of the aforementioned three surfaces is provided, where four pairs of solutions are recognized. In Figure 2, the same illustration is presented along the $z-$axis from the top. Points p_1, p_2, p_3 and p_4 denote the solution pairs that satisfy (19). Note that depending on the average power constraint, some (or all) of the points p_1, p_2, p_3 and p_4 are not admissible (for example, here p_1 is not admissible). If there is no point satisfying the average power constraint, the power allocated to the corresponding c-subchannel is zero (in order to satisfy (15)). Otherwise, there are more than one set of power allocations (P_{lr}, P_{li}) that satisfy the optimality necessary conditions. Accordingly, the power allocation could be either symmetric (corresponding to either of the points p_1, p_4) or asymmetric (corresponding to either of the points p_2, p_3). Both solution sets corresponding to points p_2 and p_3 contribute the same amount in the delivered power and transmitted information (as noted in Remark 3). Therefore, they can be chosen interchangeably.
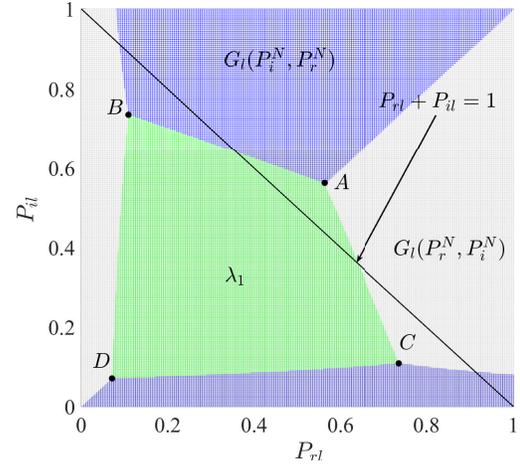
Figure 2: Illustration of Figure 1 from top view (along $z-$axis). There are 4 solutions denoted by p_1, p_2, p_3 and p_4, where point p_1 is not admissible due to contradicting the average power constraint $P_a = 1$.

Remark 4. The optimality conditions in (15) can be solved numerically using programming for solving nonlinear equations with constraints ($P_{lr}, P_{li} \geq 0$ for $l = 0, \ldots, N-1$). In Section VI, it is observed through numerical optimization that for mere WPT purposes (equivalently large values of λ_2) all the available power at the transmitter is allocated to either real or imaginary component of the strongest c-subchannel. Note that (for zero mean Gaussian inputs), although optimized for WPT, the amount of transmitted information is never zero.

VI. NUMERICAL OPTIMIZATION

In this section, we provide numerical results regarding the power allocation for different r-subchannels under a fixed average power and different delivered power constraints in order to obtain RP regions corresponding to different types of complex Gaussian inputs introduced earlier.

A. Non-zero mean inputs

We note that, the optimization problem in (10) is not convex, and accordingly, the final solution (obtained via numerical optimization) is in general a local stationary point. Due to nonconvexity of the studied problem, the final solution is dependent on the initial starting point. In order to alleviate the effect of the initial point, in our optimization, we first focus on the WPT aspect of the optimization problem with deterministic input signals[4], i.e., the variance of different r-subchannels are close to zero with a good approximation. In this case, with deterministic input signals we have $\mu_{lr} = \sqrt{P_{lr}}$, $\mu_{li} = \sqrt{P_{li}}$ for $l = 0, \cdots, N-1$. Therefore, the delivered power P_{dc} reads as

$$P_{dc} = \sum_{l=0}^{N-1} \left\{ \alpha_l |\mu_l|^4 + \left(\beta_l + g(|\mu_l|^2) \right) |\mu_l|^2 + \eta \right.$$

[4]We note that although we first optimize over deterministic signals for WPT, optimizing over means and powers for SWIPT results in the same solutions, i.e., signals with almost zero variance, however, in the expense of a long simulation time. Therefore, for the starting point of the RP region, we chose the input to be deterministic.

$$+ \Re\left\{\mu_l^2 \sum_{k=1}^{\frac{N-1}{2}} \mu_{(l+k)_N}^* \mu_{(l-k)_N}^* \Phi_{l,k}\right\}$$

$$+ \delta_{(N-1)}^l \cdot \sum_{k=1}^{\frac{N-1}{2}} \sum_{\substack{m=l+1 \\ m\neq(l+k)_N \\ m\neq(l-k)_N}}^{N-1} \Re\left\{\mu_l\mu_m\mu_{(l-k)_N}^* \mu_{(m-k)_N}^* \Psi_{l,m,k}\right\}\Bigg\}$$

$$\triangleq \sum_{l=0}^{N-1} f_{\text{WPT}}(\mu_l, h_l, N). \tag{20}$$

Accordingly, we consider the following WPT problem

$$\max_{\substack{\mu_l \\ l=0,\dots,N-1}} \sum_{l=0}^{N-1} f_{\text{WPT}}(\mu_l, h_l, N)$$

$$\text{s.t.} \quad \sum_{l=0}^{N-1} |\mu_l|^2 = P_a, \tag{21}$$

where the proof for the average power constraint satisfied with equality has been provided in Appendix B. The algorithm (WPT optimization with deterministic inputs) is run for a large number of times (here we run the algorithm 1000 times) using the Matlab command `fmincon()` with `sqp-legacy` algorithm, and each time with a new and randomly generated initial complex mean vector μ^N. After this stage, the solution corresponding to the highest delivered power P_{dc} is chosen as the initial starting point for the SWIPT optimization.

Next, in order to solve the optimization for SWIPT, we consider the following maximization, which is the weighted summation of the transmitted information and the delivered power

$$\max_{\substack{P_{lr},P_{li},\mu_{lr},\mu_{li} \\ l=0,\dots,N-1}} \sum_{l=0}^{N-1} c_0\left\{\log(1+a_l\sigma_{lr}^2) + \log(1+a_l\sigma_{li}^2)\right\}$$

$$\text{s.t.} \begin{cases} \sum_{l=0}^{N-1} P_l = P_a \\ P_{lr} \geq \mu_{lr}^2, P_{li} \geq \mu_{li}^2, \ l=0,\dots,N-1 \\ f_{ib}(P_l,\bar{P}_l,\mu_l,h_l,N) \geq P_d \end{cases}. \tag{22}$$

We solve this problem using the Matlab command `fmincon()` with `sqp-legacy` algorithm as follows. P_d is decreased gradually, starting from $P_{d,\max}$ corresponding to the maximum delivered power in (20). Decreasing P_d is terminated until it reaches $P_{d,\min}$, which is the delivered power corresponding to the water-filling solution. For the first round of the optimization, the (locally) optimal solution obtained through previous optimization (WPT with deterministic inputs) is used as the starting point (the power for different r-subchannels is considered as $P_{lr} = \mu_{lr}^2$, $P_{li} = \mu_{li}^2$, $l = 0,\dots,N-1$). Similarly, for the subsequent values of P_d, we use the solution corresponding to the previous value of P_d. The detailed description of the optimization is presented in Algorithm 1.

B. Zero mean inputs

In order to obtain the optimal power allocations for zero mean complex Gaussian inputs, we follow a similar approach

---

Algorithm 1 SWIPT algorithm (Non-zero mean inputs)

1: **procedure** WPT OPTIMIZATION
2:      $M \leftarrow$ Large number
3:      **for** $s = 1 : M$ **do**
4:          Randomly initialize μ^N, $\mu_{(s)}^{N\star} = \arg\max$ (21)
5:          $P_{dc,(s)} = \sum\limits_{l=0}^{N-1} f_{\text{WPT}}(\mu_{l,(s)}^\star, h_l, N)$, $S = \arg\max\limits_{s} P_{dc,(s)}$
6:      $P_{d,\max} \leftarrow P_{dc,(S)}$
7: **procedure** SWIPT OPTIMIZATION
8:      $P_d \leftarrow P_{d,\max}$, $s = 1$
9:      $P_{(s),r}^N \leftarrow [\mu_{0r,(S)}^{\star 2}, \dots, \mu_{(N-1)r,(S)}^{\star 2}]$
10:     $P_{(s),i}^N \leftarrow [\mu_{0i,(S)}^{\star 2}, \dots, \mu_{(N-1)i,(S)}^{\star 2}]$
11:     $\mu_{(s)}^N \leftarrow \mu_{(S)}^{N\star}$
12:     **while** $P_d > P_{d,\min}$ **do**
13:        $\{P_{(s),r}^{N\star}, P_{(s),i}^{N\star}\} = \arg\max$ (22)
14:        $\text{Inf}(s) \leftarrow \sum\limits_{l=0}^{N-1} c_0\left\{\log(1+a_l\sigma_{lr,(s)}^{\star 2}) + \log(1+a_l\sigma_{li,(s)}^{\star 2})\right\}$
15:        $P_{\text{dc}}(s) \leftarrow$ (11)
16:        $s \leftarrow (s+1)$, $P_d \leftarrow (P_d - stp)$
17:        $P_{(s),r}^N \leftarrow P_{(s-1),r}^{N\star}$, $P_{(s),i}^N \leftarrow P_{(s-1),i}^{N\star}$, $\mu_{(s)}^N \leftarrow \mu_{(s-1)}^{N\star}$

---

presented in Section VI-A. The optimization problem considered here is given as

$$\max_{\substack{P_{lr},P_{li} \\ l=0,\dots,N-1}} \sum_{l=0}^{N-1} c_0\left\{\log(1+a_lP_{lr}) + \log(1+a_lP_{li})\right\}$$

$$\text{s.t.} \begin{cases} \sum_{l=0}^{N-1} P_l = P_a \\ P_{lr} \geq 0, P_{li} \geq 0, \ l=0,\dots,N-1 \\ f_{ib}(P_l,\bar{P}_l,0,h_l,N) = P_d \end{cases}. \tag{23}$$

The optimization is explained in Algorithm 2[5].

C. Numerical results

In this section, we present the results obtained through numerical simulations. First, we focus on the optimized RP regions corresponding to different types of channel inputs. Later, we compare the constellation of optimized non-zero mean and zero mean complex Gaussian inputs on different points of their corresponding optimized RP regions.

In Figure 3, the RP regions for *Asymmetric Non-zero mean Gaussian* (ANG), studied in this paper and *Symmetric Non-zero mean Gaussian* (SNG) studied in [8] and *Zero mean Gaussian*

---

[5]We note that, as an alternative approach, the optimality conditions in Lemma 1 can be used in order to find the optimal power allocations. To do so, solving the nonlinear equations (14) and (15) have to be considered with the constraints $P_{li}, P_{li} \geq 0$. Accordingly, one can use the MATLAB command `fsolve()`. The optimization is initialized with a very small (in norm) power vector and each time the vector is updated until a condition on convergence is met.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2927221, IEEE Transactions on Communications

8

Algorithm 2 SWIPT algorithm (Zero mean inputs)

1: **procedure** SWIPT OPTIMIZATION
2: $\quad P_d \leftarrow P_{d,\max}$, $M \leftarrow$ Large number
3: $\quad$ **while** $P_d > P_{d,\min}$ **do**
4: $\quad\quad$ **for** $t = 1 : M$ **do**
5: $\quad\quad\quad$ Randomly initialize $\{P_r^N, P_i^N\}$
6: $\quad\quad\quad$ $\{P_{(t),r}^{N\star}, P_{(t),i}^{N\star}\} = \arg\max$ (23)
7: $\quad\quad\quad$ $IP(t) =$ Calculate cost in (23) for $\{P_{(t),r}^{N\star}, P_{(t),i}^{N\star}\}$
8: $\quad\quad$ $S = \arg\max\limits_{t} IP(t)$
9: $\quad\quad$ save $\{P_{(S),r}^{N\star}, P_{(S),i}^{N\star}\}$
10: $\quad\quad$ $P_d \leftarrow (P_d - stp)$

Figure 3: The optimized RP regions corresponding to ANG, SNG, ZG and ZGL with an average power constraint $P_a = 100~\mu$W and SNR 20 dB.

(ZG) are illustrated[6]. We also obtain the RP region corresponding to the optimal power allocations under the linear model assumption of the EH. This is done by obtaining the power allocations from [3, Equation (9)] for different constraints and calculating the corresponding delivered power and transmitted information. This region is denoted by *Zero mean Gaussian for Linear model* (ZGL). As it is observed in Figure 3, due to the asymmetric power allocation in ANG, there is an improvement in the RP region compared to SNG. Additionally, it is observed that ANG and SNG achieve larger RP region compared to optimized ZG and that ANG and SNG perform better than ZGL (highlighting the fact that for scenarios that the nonlinear model for EH is valid, ZGL is not optimal any more). The main reason of improvement in the RP regions corresponding to ANG, SNG is due to the fact that allowing the mean of the channel inputs to be non-zero boosts the fourth order term (more explanations can be found in [8]) in (6), resulting in more contribution in the delivered power at the receiver.

In Figure 4, from left to right, the optimized inputs in terms of their complex mean μ_l, $l = 0,\ldots,8$ (represented as dots) and their corresponding r-subchannel variances $\sigma_{lr}^2, \sigma_{li}^2$, $l = 0,\ldots,8$ (represented as ellipses) are shown for points A, B and C in Figure 3, respectively. Point A represents the maximum delivered power with the zero transmitted information (note that information of a deterministic signal is zero). Point B represents the performance of a typical input used for power and information transmission. Finally, point C represents the performance of an input obtained via water-filling (the delivered power constraint is inactive). From these 3 plots it is observed that as we move from point A to point C, the mean of different r-subchannels decrease, however, they (means of different r-subchannels) keep their relative structure, roughly. Also, as we move to point C, the means of different r-subchannels get to zero with their variances increasing asymmetrically until the power allocation gets to water-filling solution (where the

power allocation between the real and imaginary components are symmetric). This result is in contrast with the results in [8], where the power allocation to the real and imaginary components in each c-subchannel is symmetric. Similar results regarding the benefit of asymmetric power allocation has also been reported in [16] for deterministic AWGN channel with nonlinear EH.

In Figure 3, the point D corresponds to the input, where all of the c-subchannels other than the strongest one (in terms of the $\max\limits_{l=0,\ldots,N-1} |h_l|^2$) are with zero power. For the strongest c-subchannel, at point D, all the transmit power is allocated to either real or imaginary component of the c-subchannel. The reason for this observation is explained in Remark 4. This observation is also inline with the result of [16], where it is shown that for a flat fading channel, the maximum power is obtained by allocating all the transmitter power to only one r-subchannel. Note that this is different from the power allocation under the linear model assumptio for the EH (i.e. ZGL), for which all the transmit power is allocated to the strongest c-subchannel (equally divided between the ream and imaginary components of the input) to maximize delivered power.

In Figure 5, the variances of different r-subchannels corresponding to the point E in Figure 3 are illustrated. Numerical optimization reveals that, as we move from point D to point C (increasing the information demand at the receiver) in Figure 3, the variance of the strongest c-subchannel varies asymmetrically (in its real and imaginary components). This observation can be justified as follows. For higher delivered power demands, the strongest c-subchannel receives a power allocation similar to the solutions p_2 or p_3 in Figure 2, whereas the other c-subchannels take the power allocation corresponding to the point p_4 in Figure 2[7]. The power allocation in point C is the water-filling solution. .

---

[6]In our numerical results, we have used the channel in [8, Figure 3]. For a given number of c-subchannels N, the amplitude of each c-subchannel is chosen based on the channel in [8, Figure 3] by dividing the channel into equal sub-bands. The phase of the c-subchannel is chosen arbitrarily based on a uniform distribution over $[0, 2\pi]$. For example for the simulation results in Figure 3 we have $N = 9$ c-subchannels with coefficients as $[-1.79+0.14i, -0.59-1.94i, -0.14-2.09i, 0.89-1.85i, -1.92-0.14i, -1.76+0.29i, -1.61 - 0.05i, 1.37 + 0.29i, 1.27 + 0.22i]$.

[7]For very low average power constraints, it is observed that the power allocation is symmetric across all the c-subchannels. This can be justified by noting that for very low average power constraints, the admissible power allocations correspond to solutions similar to the point D in Figure 2.

Figure 5: Representation of variances of different c-subchannels corresponding to the point E in Figure 3. The strongest c-subchannel receives more power compared to the other c-subchannels and the other c-subchannels attain CSCG inputs. The axis values are with the unit mVolt.
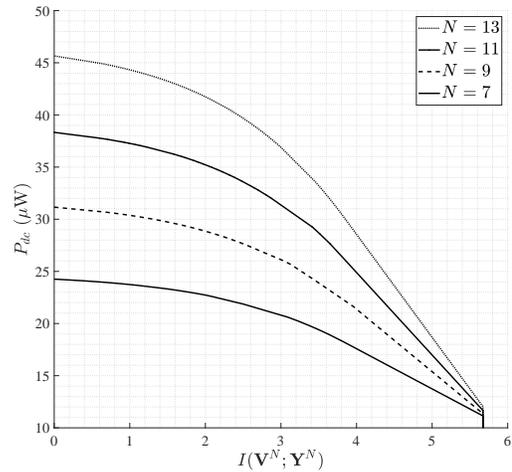
Figure 6: The optimized RP regions corresponding to nonezero mean Gaussian inputs for $N = 3, 5, 7, 9$ with an average power constraint $P_a = 100$ μW and SNR 20 dB.

in the delivered power in (11), and is inline with observations made in [8], [10][8].

As another interesting observation, in Figure 7, the numerically optimized inputs for WPT under the assumption of flat fading for the channel are illustrated for $N = 3, 5, 7, 9$. In Figure 7, the optimized inputs for WPT purposes (zero variance inputs) are illustrated. It is observed that the phases of the mean on different c-subchannels are also equally spaced.

Figure 4: From top to bottom, the mean and variance of different r-subchannels' input corresponding to the points A, B and C in Figure 3, respectively. As we move forward from point A to point C, the variance of different r-subchannels increase, whereas their corresponding means shrink to zero. The axis values are with the unit mVolt.

Remark 5. *In Figure 6 (using the optimization algorithm, explained earlier in Algorithm 1) the RP regions are obtained for $N = 7, 9, 11, 13$. To make the results comparable we assume the the channel is flat fading with $h_l = 0.5 + j0.5$ for $l = 1, \ldots, N$. It is observed that the delivered power at the receiver is increased by the number of the c-subchannels N. This is due to the presence of input moments (higher than 2)*

---

[8]We note that, in practical implementations, this observation (increasing delivered power with N) cannot be valid for all N, and the delivered power is saturated after some N. This is due to the diode breakdown effect, which has not been considered in our model (6) due to small signal analysis. This is further discussed in [8].
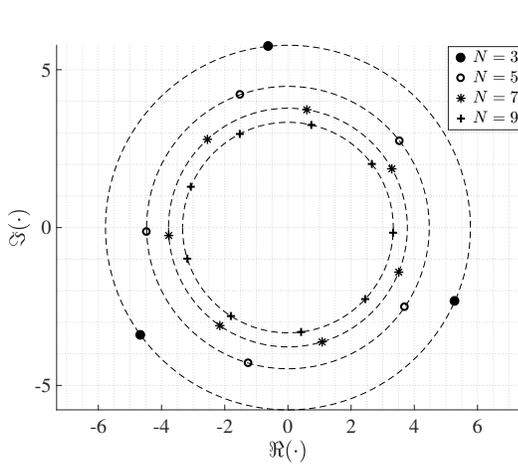
Figure 7: The numerically optimized inputs for WPT (under flat fading assumption for the channel) for $N = 3, 5, 7, 9$ with an average power constraint $P_a = 100\ \mu$W and SNR 20 dB. The axis values are with the unit mVolt.

VII. CONCLUSION

In this paper, we studied SWIPT signalling for frequency-selective channels under transmit average power and receiver delivered power constraints. We considered an approximation for the nonlinear EH, which is based on truncation (up to fourth moment) of the Taylor expansion of the rectenna's diode characteristic function. For independent input distributions on different r-subchannels and i.i.d. inputs on each r-subchannel, we obtained the delivered power in terms of the system baseband parameters, which demonstrates the dependency of the delivered power on the mean as well as higher moments of the channel input. Assuming that the channel input is constrained to Gaussian distributions, numerical observations reveal that in general non-zero mean Gaussian inputs attain a larger RP region compared to their zero mean counterparts. As a special scenario, for zero mean Gaussian inputs, we obtained the conditions for optimal power allocation on different r-subchannels. Using numerical optimization, it is observed that optimized non-zero mean inputs (with asymmetric power allocation in each c-subchannel) achieve larger RP region, compared to their optimized zero mean as well as non-zero mean (with symmetric power allocation in each c-subchannel [8]) counterparts.

A. PROOF OF THE PROPOSITION 1

In the following, we obtain the baseband equivalent of (6). Considering first the term $\mathbb{E}\mathcal{E}[Y_{\mathrm{rf}}(t)^2]$, we have

$$
\begin{aligned}
\mathbb{E}\mathcal{E}[Y_{\mathrm{rf}}(t)^2] &= \frac{1}{2}\mathbb{E}\mathcal{E}\left[\left(Y(t)e^{jf_c t} + Y^*(t)e^{-jf_c t}\right)^2\right] \\
&= \mathbb{E}\mathcal{E}\left[|Y(t)|^2\right] \\
&= \mathbb{E}\mathcal{E}\left[\sum_{n,m} Y[n]Y[m]^*\mathrm{sinc}(f_w t - n)\mathrm{sinc}(f_w t - m)\right] \\
&= \sum_{n,m} \mathbb{E}\left[Y[n]Y[m]^*\right]\mathcal{E}\left[\mathrm{sinc}(f_w t - n)\mathrm{sinc}(f_w t - m)\right]
\end{aligned}
\tag{24}
$$

$$
= \lim_{T\to\infty} \frac{1}{Tf_w}\left(\sum_n \mathbb{E}\left[|Y[n]|^2\right]\right)
\tag{25}
$$

$$
= \lim_{T\to\infty} \frac{1}{Tf_w}\left(\sum_{s=1}^{N_b}\sum_{n=(s-1)(L+N-1)}^{s(L+N-1)} \mathbb{E}\left[|Y[n]|^2\right]\right)
$$

$$
\simeq \lim_{T\to\infty} \frac{1}{Tf_w}\left(\sum_{s=1}^{N_b}\sum_{n=(s-1)(L+N-1)+L}^{s(L+N-1)} \mathbb{E}\left[|Y[n]|^2\right]\right)
\tag{26}
$$

$$
= \lim_{T\to\infty} \frac{1}{Tf_w}\left(\sum_{s=1}^{N_b}\sum_{l=(s-1)N}^{sN-1} \mathbb{E}\left[|\mathbf{y}_l|^2\right]\right)
\tag{27}
$$

$$
= \lim_{T\to\infty} \frac{1}{Tf_w}\sum_{s=1}^{N_b}\sum_{l=(s-1)N}^{sN-1}\left(|h_l|^2 P_l + \sigma_w^2\right)
$$

$$
= \sum_{l=0}^{N-1}\left(|h_l|^2 P_l + \sigma_w^2\right),
\tag{28}
$$

where (24) is due to $\mathcal{E}[Y(t)^2 e^{2jf_c t}] = \mathcal{E}[Y(t)^{*2} e^{-2jf_c t}] = 0$. (25) is due to

$$
\int_{-\infty}^{\infty} \mathrm{sinc}(f_w t - n)\mathrm{sinc}(f_w t - m)dt = \frac{1}{f_w}\delta(m - n).
$$

In (26), s is the OFDM symbol index and $N_b = f_w T/(N + L)$. Note that, in (26) we neglect the delivered power related to the cyclic prefix (CP). This is mainly due to simplifying the expressions and in practice the EH can capture the power of the CP in the same way as for the other parts of the signal. In (27) the result is due to Parseval's theorem.

Next, considering the second term in (6), i.e., $\mathbb{E}\mathcal{E}[Y_{\mathrm{rf}}(t)^4]$, we have

$$
\begin{aligned}
\mathbb{E}\mathcal{E}[Y_{\mathrm{rf}}(t)^4] &= \frac{1}{4}\mathbb{E}\mathcal{E}\left[4|Y(t)|^4 + (Y(t)^2 e^{j2f_c t} + Y(t)^{*2}e^{-j2f_c t})^2\right. \\
&\qquad \left. +4|Y(t)|^2(Y(t)^2 e^{j2f_c t} + Y(t)^{*2}e^{-j2f_c t})\right] \\
&= \frac{3}{2}\mathbb{E}\mathcal{E}\left[|Y(t)|^4\right].
\end{aligned}
\tag{29}
$$

Note that the signal $|Y(t)|^2$ is real with bandwidth $(-f_w, f_w)$. Therefore, (29) can be rewritten as

$$
\frac{3}{2}\mathbb{E}\mathcal{E}\left[|Y(t)|^4\right] = \lim_{T\to\infty}\frac{3}{4Tf_w}\sum_n\left(\mathbb{E}\left[|Y[n]|^4\right] + \mathbb{E}\left[|\tilde{Y}[n]|^4\right]\right)
\tag{30}
$$

$$
= \lim_{T\to\infty}\frac{3}{4Tf_w}\sum_{s=1}^{N_b}\sum_{n=(s-1)(L+N-1)+L}^{s(L+N-1)}\left(\mathbb{E}\left[|Y[n]|^4\right] + \mathbb{E}\left[|\tilde{Y}[n]|^4\right]\right)
$$

$$
\simeq \frac{3}{4}\sum_{n=L}^{N+L-1}\left(\mathbb{E}\left[|Y[n]|^4\right] + \mathbb{E}\left[|\tilde{Y}[n]|^4\right]\right),
$$

where $Y[n]$ and $\tilde{Y}[n]$ in (30) are the samples of the process $Y(t)$ taken at times $t = \frac{2n}{2f_w}$ and $t = \frac{2n+1}{2f_w}$, respectively, i.e., $Y[n] \triangleq Y(\frac{2n}{2f_w})$ and $\tilde{Y}[n] \triangleq Y(\frac{2n+1}{2f_w})$. In the following, we analyze $Y[n]$ and $\tilde{Y}[n]$, separately.

A. Samples of $Y(t)$ at times $t = \frac{2n}{2f_w}$:

First considering $Y[n]$, (in one OFDM symbol) we have

$$
\sum_{n=L}^{L+N-1} \mathbb{E}\left[|Y[n]|^4\right] = \sum_{n=L}^{L+N-1} \mathbb{E}\left[(|Y[n]|^2)^2\right]
$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}\left[|\mathbf{y}_k \otimes \mathbf{y}^*_{(-k)_N}|^2 \right] \tag{31}$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} \mathbb{E}\left[\mathbf{y}_l \mathbf{y}^*_{(l-k)_N} \mathbf{y}^*_m \mathbf{y}_{(m-k)_N} \right], \tag{32}$$

where (31) is due to Parseval's theorem and convolution property of DFT. Define $\mathbf{u}_l \triangleq h_l \mathbf{v}_l$. We have $\mathbf{y}_l = \mathbf{u}_l + \mathbf{w}_l$. By expanding (32) we have

$$
\begin{aligned}
&\mathbb{E}\left[\mathbf{y}_l \mathbf{y}^*_{(l-k)_N} \mathbf{y}^*_m \mathbf{y}_{(m-k)_N} \right] \\
&= \mathbb{E}[\mathbf{u}_l \mathbf{u}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{u}_l \mathbf{u}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{w}_{(m-k)_N} \\
&\quad + \mathbf{u}_l \mathbf{u}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{u}_l \mathbf{u}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{w}_{(m-k)_N} \\
&\quad + \mathbf{u}_l \mathbf{w}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{u}_l \mathbf{w}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{w}_{(m-k)_N} \\
&\quad + \mathbf{u}_l \mathbf{w}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{u}_l \mathbf{w}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{w}_{(m-k)_N} \\
&\quad + \mathbf{w}_l \mathbf{u}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{w}_l \mathbf{u}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{w}_{(m-k)_N} \\
&\quad + \mathbf{w}_l \mathbf{u}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{w}_l \mathbf{u}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{w}_{(m-k)_N} \\
&\quad + \mathbf{w}_l \mathbf{w}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{w}_l \mathbf{w}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{w}_{(m-k)_N} \\
&\quad + \mathbf{w}_l \mathbf{w}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{w}_l \mathbf{w}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{w}_{(m-k)_N}] \tag{33}
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbb{E}[\mathbf{u}_l \mathbf{u}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{u}_l \mathbf{u}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{w}_{(m-k)_N} \\
&\quad + \mathbf{u}_l \mathbf{w}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{w}_{(m-k)_N} + \mathbf{w}_l \mathbf{u}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{u}_{(m-k)_N} \\
&\quad + \mathbf{w}_l \mathbf{w}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{u}_{(m-k)_N} + \mathbf{w}_l \mathbf{w}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{w}_{(m-k)_N}]. \tag{34}
\end{aligned}
$$

In the following, we calculate each of the terms in (34) (Note that since the noise is CSCG, we have $\mathbb{E}[|\mathbf{w}|^4] = 2\sigma_w^4$ and $\mathbb{E}[\mathbf{w}^{*2}] = \mathbb{E}[\mathbf{w}^2] = 0$)

$$
\begin{aligned}
&\sum_{k,l,m=0}^{N-1} \mathbb{E}\left[\mathbf{w}_l \mathbf{w}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{w}_{(m-k)_N} \right] \\
&= \sum_{k,l,m=0}^{N-1} \delta_{(l-m)_N}(\delta_k 2\sigma_w^4 + (1-\delta_k)\sigma_w^4) + (1 - \delta_{(l-m)_N})\delta_k \sigma_w^4 \\
&= \sum_{l=0}^{N-1} 2\sigma_w^4 + \sum_{l=0}^{N-1}\sum_{k=1}^{N-1} \sigma_w^4 + \sum_{l=0}^{N-1}\sum_{m=0,m\neq l}^{N-1} \sigma_w^4 \\
&= \sum_{l=0}^{N-1} 2N^2 \sigma_w^4, \tag{35}
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{k,l,m=0}^{N-1} \mathbb{E}\left[\mathbf{w}_l \mathbf{w}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{u}_{(m-k)_N} \right] \\
&= \sum_{k,l,m=0}^{N-1} \delta_k \sigma_w^2 |h_m|^2 \mathbb{E}[|V_m|^2] \\
&= \sum_{l=0}^{N-1} N\sigma_w^2 |h_l|^2 P_l, \tag{36}
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{k,l,m=0}^{N-1} \mathbb{E}\left[\mathbf{w}_l \mathbf{u}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{u}_{(m-k)_N} \right] \\
&= \sum_{k,l,m=0}^{N-1} \delta_{(l-m)_N} \sigma_w^2 |h_{(m-k)_N}|^2 \mathbb{E}\left[|\mathbf{v}_{(m-k)_N}|^2 \right] \\
&= \sum_{k,m} \sigma_w^2 |h_{(m-k)_N}|^2 \mathbb{E}\left[|\mathbf{v}_{(m-k)_N}|^2 \right]
\end{aligned}
$$

$$= \sum_{l=0}^{N-1} N\sigma_w^2 |h_l|^2 P_l, \tag{37}$$

where (37) is due to the property of circular convolution. For the other terms we have

$$
\begin{aligned}
&\sum_{k,l,m=0}^{N-1} \mathbb{E}\left[\mathbf{u}_l \mathbf{w}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{w}_{(m-k)_N} \right] \\
&= \sum_{k,l,m=0}^{N-1} \delta_{(l-m)_N} \sigma_w^2 |h_m|^2 P_m \\
&= \sum_{l=0}^{N-1} N\sigma_w^2 |h_l|^2 P_l, \tag{38}
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{k,l,m=0}^{N-1} \mathbb{E}\left[\mathbf{u}_l \mathbf{u}^*_{(l-k)_N} \mathbf{w}^*_m \mathbf{w}_{(m-k)_N} \right] \\
&= \sum_{k,l,m=0}^{N-1} \delta_k \sigma_w^2 |h_l|^2 \mathbb{E}[|\mathbf{v}_l|^2] \\
&= \sum_{l=0}^{N-1} N\sigma_w^2 |h_l|^2 P_l, \tag{39}
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{k,l,m=0}^{N-1} \mathbb{E}\left[\mathbf{u}_l \mathbf{u}^*_{(l-k)_N} \mathbf{u}^*_m \mathbf{u}_{(m-k)_N} \right] \\
&= \sum_{l,m,k=0}^{N-1} \Big\{ (1 - \delta_{(l-m)_N})(1 - \delta_k) E_{l,m,k} \\
&\quad + \delta_{(l-m)_N}(1 - \delta_k)|h_l|^2 |h_{(l-k)_N}|^2 P_l P_{(l-k)_N} \\
&\quad + \delta_{(l-m)_N} \delta_k |h_l|^4 Q_l + (1 - \delta_{(l-m)_N})\delta_k |h_l|^2 |h_m|^2 P_l P_m \Big\} \\
&= \sum_{l,m,k=0}^{N-1} (1 - \delta_{(l-m)_N})(1 - \delta_k) E_{l,m,k} \\
&\quad + \sum_{k=1}^{N-1}\sum_{l=0}^{N-1} |h_l|^2 |h_{(l-k)_N}|^2 P_l P_{(l-k)_N} \\
&\quad + \sum_{l=0}^{N-1} |h_l|^4 Q_l + \sum_{l=0}^{N-1}\sum_{m=0,m\neq l}^{N-1} |h_l|^2 |h_m|^2 P_l P_m \\
&= \sum_{l,m,k=0}^{N-1} (1 - \delta_{(l-m)_N})(1 - \delta_k) E_{l,m,k} \\
&\quad + 2\sum_{l=0}^{N-2}\sum_{m=l+1}^{N-1} |h_l|^2 |h_m|^2 P_l P_m \\
&\quad + \sum_{l=0}^{N-1} |h_l|^4 Q_l + 2\sum_{l=0}^{N-2}\sum_{m=l+1}^{N-1} |h_l|^2 |h_m|^2 P_l P_m \\
&= \sum_{l,m,k=0}^{N-1} (1 - \delta_{(l-m)_N})(1 - \delta_k) E_{l,m,k} \\
&\quad + 4\sum_{l=0}^{N-1} \delta_{N-1}^l \sum_{m=l+1}^{N-1} |h_l|^2 |h_m|^2 P_l P_m + \sum_{l=0}^{N-1} |h_l|^4 Q_l. \tag{40}
\end{aligned}
$$

For $E_{l,m,k}$ we have $l \neq m$, $k \neq 0$. Hence, we also have $l \neq (l - k)_N$, $m \neq (m - k)_N$. Four different situations occur. We

have

$$E_{l,m,k}$$

$$
= \begin{cases}
\mathbb{E}[\boldsymbol{u}_l^2]\mathbb{E}[\boldsymbol{u}_m^{*\,2}] & l=(m-k)_N, m=(l-k)_N \\
\mathbb{E}[\boldsymbol{u}_l^2]\mathbb{E}[\boldsymbol{u}_m^*]\mathbb{E}[\boldsymbol{u}_{(l-k)_N}^*] & l=(m-k)_N, m\neq(l-k)_N \\
\mathbb{E}[\boldsymbol{u}_m^{*\,2}]\mathbb{E}[\boldsymbol{u}_l]\mathbb{E}[\boldsymbol{u}_{(m-k)_N}] & l\neq(m-k)_N, m=(l-k)_N \\
\mathbb{E}[\boldsymbol{u}_l]\mathbb{E}[\boldsymbol{u}_m^*]\mathbb{E}[\boldsymbol{u}_{(l-k)_N}^*]\mathbb{E}[\boldsymbol{u}_{(m-k)_N}] & l\neq(m-k)_N, m\neq(l-k)_N
\end{cases}
$$

$$
= \begin{cases}
h_l^2 h_m^{*\,2}\bar{P}_l\bar{P}_m^* & l=(m-k)_N, m=(l-k)_N \\
h_l^2 h_m^* h_{(l-k)_N}^*\bar{P}_l\mu_m^*\mu_{(l-k)_N}^* & l=(m-k)_N, m\neq(l-k)_N \\
h_m^{*2}h_l h_{(m-k)_N}\bar{P}_m^*\mu_l\mu_{(m-k)_N} & l\neq(m-k)_N, m=(l-k)_N \\
a_l a_m^* a_{(l-k)_N}^* a_{(m-k)_N} & l\neq(m-k)_N, m\neq(l-k)_N
\end{cases},
$$

(41)

where $a_l \triangleq h_l\mu_l$. Substituting the terms in (41) in $E_{l,m,k}$ in (40), we have

$$
\sum_{k,l,m=0}^{N-1}\mathbb{E}\left[\boldsymbol{u}_l\boldsymbol{u}_{(l-k)_N}^*\boldsymbol{u}_m^*\boldsymbol{u}_{(m-k)_N}\right]
$$

$$
= \sum_{\substack{k\neq0,l\neq m\\l=(m-k)_N\\m=(l-k)_N}} h_l^2 h_m^{*\,2}\bar{P}_l\bar{P}_m^* + \sum_{\substack{k\neq0,l\neq m\\l=(m-k)_N\\m\neq(l-k)_N}} h_l^2 h_m^* h_{(l-k)_N}^*\bar{P}_l\mu_m^*\mu_{(l-k)_N}^*
$$

$$
+ \sum_{\substack{k\neq0,l\neq m\\l\neq(m-k)_N\\m=(l-k)_N}} h_m^{*2}h_l h_{(m-k)_N}\bar{P}_m^*\mu_l\mu_{(m-k)_N}
$$

$$
+ \sum_{\substack{k\neq0,l\neq m\\l\neq(m-k)_N\\m\neq(l-k)_N}} h_l h_m^* h_{(l-k)_N}^* h_{(m-k)_N}\mu_l\mu_m^*\mu_{(l-k)_N}^*\mu_{(m-k)_N}
$$

$$
+ 4\sum_{l=0}^{N-1}\delta_{N-1}^l\sum_{m=l+1}^{N-1}|h_l|^2|h_m|^2 P_l P_m + \sum_{l=0}^{N-1}|h_l|^4 Q_l. \quad (42)
$$

For the first term in the RHS of (42), it is verified that (we recall that N is odd)

$$
\sum_{\substack{k\neq0,l\neq m\\l=(m-k)_N\\m=(l-k)_N}} h_l^2 h_m^{*\,2}\bar{P}_l\bar{P}_m^* = 0. \quad (43)
$$

This is because from $l=(m-k)_N$ and $m=(l-k)_N$, we have $l=(l-2k)_N$, which results in $k=0,N/2$. Noting that N is odd and $k\neq0$, there is no such integer. For the second and third terms in the RHS of (42), we have $l=(m-k)_N$ and $m=(l-k)_N$, respectively. Therefore, we obtain

$$
\sum_{\substack{k\neq0,l\neq m\\l=(m-k)_N\\m\neq(l-k)_N}} h_l^2\bar{P}_l h_m^* h_{(l-k)_N}^*\mu_m^*\mu_{(l-k)_N}^*
$$

$$
= \sum_{l=0}^{N-1}\sum_{k=1}^{N-1} h_l^2\bar{P}_l h_{(l+k)_N}^* h_{(l-k)_N}^*\mu_{(l+k)_N}^*\mu_{(l-k)_N}^*, \quad (44)
$$

$$
\sum_{\substack{k\neq0,l\neq m\\l\neq(m-k)_N\\m=(l-k)_N}} h_m^{*2}\bar{P}_m^* h_l h_{(m-k)_N}\mu_l\mu_{(m-k)_N}
$$

$$
= \sum_{l=0}^{N-1}\sum_{k=1}^{N-1} h_l^{*2}\bar{P}_l^* h_{(l+k)_N} h_{(l-k)_N}\mu_{(l+k)_N}\mu_{(l-k)_N}. \quad (45)
$$
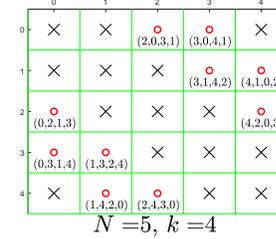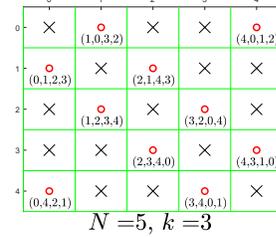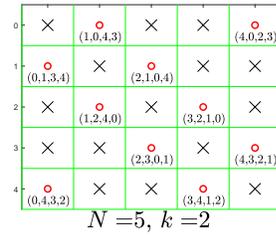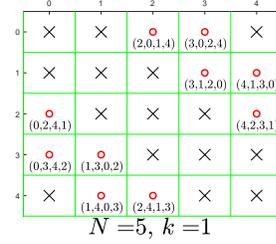
Figure A1: Illustration of the admissible indices for the fourth term in the RHS of (42) for $N = 5$. The red dots are the admissible triplet indices (l,m,k), whereas the others are not. The quadruplet indices under each red dot represents the numbers $(l,m,(l-k)_N,(m-k)_N)$ corresponding to the indices of each term for the fourth summation in the RHS of (42). Note the diagonal symmetry for a given k, as well as the symmetry between ($k=1$ and $k=4$) or ($k=2$ and $k=3$).

Since the terms in (44) and (45) are conjugate of each other, therefore by adding the two, we get

Term in (44) + Term in (45)

$$
= 2\sum_{l=0}^{N-1}\sum_{k=1}^{N-1}\Re\{h_l^2\bar{P}_l h_{(l+k)_N}^* h_{(l-k)_N}^*\mu_{(l+k)_N}^*\mu_{(l-k)_N}^*\}. \quad (46)
$$

For the fourth term in the RHS of (42), note that since $k\neq0$, $m\neq l$, $l\neq(m-k)_N$ and $m\neq(l-k)_N$, overall we have $N(N-1)(N-3)$ terms. Before simplifying this term, we mention the following two points[9]:

[9]We recall that N is odd. Similar discussion can be used to simplify the results for even N. However, since calculations for odd N is easier to follow, we opt to bring the discussion for odd N only.

- **Symmetry with respect to the index** k: Note that for the fourth term in the RHS of (42) for $k = i$, $i = 1, \ldots, \frac{N-1}{2}$, if the term $a_l a_m^* a_{(l-i)_N}^* a_{(m-i)_N}$ [10] is admissible (the red circles in Figure A1), then it is easy to verify that there exists the same term in the ordering of $a_{(m-i)_N} a_{(l-i)_N}^* a_m^* a_l$ [11] with the corresponding index $k = N - i$. As an example for clarification, this can also be verified by considering the tables corresponding to $k = 1$ and $k = 4$ (or $k = 2$ and $k = 3$) together in Figure A1. Therefore, due to this property it is enough to consider the summation over the indices $k = 1 : \frac{N-1}{2}$ multiplied by a factor two.

- **Diagonal symmetry between m and l for a given** k: Note that for a fixed index k, the admissible term $a_l a_m^* a_{(l-k)_N}^* a_{(m-k)_N}$ is conjugate of the admissible term $a_m a_l^* a_{(m-k)_N}^* a_{(l-k)_N}$. This accordingly shows in the table of indices corresponding to the given index k (similar to each one of the tables in Figure A1), that the upper-diagonal admissible set of elements are conjugate of the lower-diagonal admissible ones. Since we consider the summation of these elements, it is enough to run the summation over the twice of real of lower-diagonal set of admissible elements.

Therefore, according to the aforementioned points, the fourth term on the RHS of (42) is simplified as

$$\sum_{\substack{k \neq 0, l \neq m \\ l \neq (m-k)_N \\ m \neq (l-k)_N}} a_l a_m^* a_{(l-k)_N}^* a_{(m-k)_N}$$

$$= 4 \sum_{k=1}^{\frac{N-1}{2}} \sum_{l=0}^{N-2} \sum_{\substack{m=l+1 \\ m \neq (l+k)_N \\ m \neq (l-k)_N}}^{N-1} \Re\{a_l a_m^* a_{(l-k)_N}^* a_{(m-k)_N}\} \quad (47)$$

Substituting (43), (46) and (47) in (42), we have

$$\sum_{k,l,m} \mathbb{E}\left[u_l u_{(l-k)_N}^* u_m^* u_{(m-k)_N} \right]$$

$$= 4 \sum_{l=0}^{N-1} \delta_{N-1}^l \sum_{m=l+1}^{N-1} |h_l|^2 |h_m|^2 P_l P_m + \sum_{l=0}^{N-1} |h_l|^4 Q_l \quad (48)$$

$$+ 2 \sum_{l=0}^{N-1} \sum_{k=1}^{N-1} \Re\{h_l^2 \bar{P}_l h_{(l+k)_N}^* h_{(l-k)_N}^* \mu_{(l+k)_N}^* \mu_{(l-k)_N}^*\}$$

$$+ 4 \sum_{l=0}^{N-2} \sum_{k=1}^{\frac{N-1}{2}} \sum_{\substack{m=l+1 \\ m \neq (l+k)_N \\ m \neq (l-k)_N}}^{N-1} \Re\{h_l h_m^* h_{(l-k)_N}^* h_{(m-k)_N}$$

$$\cdot \mu_l \mu_m^* \mu_{(l-k)_N}^* \mu_{(m-k)_N}\} \quad (49)$$

B. Samples of $Y(t)$ at times $t = \frac{2n+1}{2f_w}$:

The continuous baseband received signal $Y(t)$ can be written as [24, Chapter 2]

$$Y(t) = \sum_k X[k] \sum_i a_i^b(t) \mathrm{sinc}(f_w t - f_w \tau_i(t) - k) + W(t). \quad (50)$$

Considering the samples at times $t = \frac{2n+1}{f_w}$ for integer n, we have

$$Y\left(\frac{2n+1}{2f_w}\right)$$

$$= \sum_k X[k] \sum_i a_i^b\left(\frac{2n+1}{2f_w}\right) \mathrm{sinc}\left(\frac{2n+1}{2} - f_w \tau_i\left(\frac{2n+1}{2f_w}\right) - k\right)$$

$$+ W\left(\frac{2f_w+1}{2f_w}\right) \quad (51)$$

$$= \sum_k X[k] \sum_i a_i^b\left(\frac{2n+1}{2f_w}\right) \mathrm{sinc}\left(n + \tfrac{1}{2} - f_w \tau_i\left(\frac{2n+1}{2f_w}\right) - k\right)$$

$$+ W\left(\frac{n}{f_w} + \frac{1}{2f_w}\right) \quad (52)$$

$$= \sum_l X[n-l] \sum_i a_i^b\left(\frac{2n+1}{2f_w}\right) \mathrm{sinc}\left(l + \tfrac{1}{2} - f_w \tau_i\left(\frac{2n+1}{2f_w}\right)\right)$$

$$+ W\left(\frac{n}{f_w} + \frac{1}{2f_w}\right) \quad (53)$$

$$= \sum_{l=0}^{L-1} X[n-l] \tilde{h}^u[l] + W\left(\frac{n}{f_w} + \frac{1}{2f_w}\right), \quad (54)$$

where $\tilde{h}_l^u \triangleq \sum_i a_i^b\left(\frac{2n+1}{2f_w}\right) \mathrm{sinc}\left(l + \tfrac{1}{2} - f_w \tau_i\left(\frac{2n+1}{2f_w}\right)\right)$ stands for the samples (in the time domain) of the channel at times $t = \frac{2n+1}{2f_w}$. Recalling that $\tilde{Y}[n] \triangleq Y\left(\frac{2n+1}{2f_w}\right)$, we have

$$\sum_{n=L}^{L+N-1} \mathbb{E}\left[\left|\tilde{Y}[n]\right|^4\right]$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}\left[|\tilde{y}_k \otimes \tilde{y}_{-k}^*|^2\right] \quad (55)$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} \mathbb{E}\left[\tilde{y}_l \tilde{y}_{(l-k)_N}^* \tilde{y}_m^* \tilde{y}_{(m-k)_N}\right]. \quad (56)$$

Similarly to (33), we have [12]

$$\mathbb{E}\left[\tilde{y}_l \tilde{y}_{(l-k)_N}^* \tilde{y}_m^* \tilde{y}_{(m-k)_N}\right]$$

$$= \mathbb{E}[u_l u_{(l-k)_N}^* u_m^* u_{(m-k)_N} + u_l u_{(l-k)_N}^* w_m^* w_{(m-k)_N}$$

$$+ u_l w_{(l-k)_N}^* u_m^* w_{(m-k)_N} + w_l u_{(l-k)_N}^* w_m^* u_{(m-k)_N}$$

$$+ w_l w_{(l-k)_N}^* u_m^* u_{(m-k)_N} + w_l w_{(l-k)_N}^* w_m^* w_{(m-k)_N}]. \quad (57)$$

Following the same steps in (35), (36), (37), (38), (39) and (49), for each term in (57) we have

$$\sum_{k,l,m} \mathbb{E}\left[w_l w_{(l-k)_N}^* w_m^* w_{(m-k)_N}\right] = \sum_{l=0}^{N-1} 2N^2 \sigma_w^4, \quad (58)$$

$$\sum_{k,l,m} \mathbb{E}\left[w_l w_{(l-k)_N}^* u_m^* u_{(m-k)_N}\right] = \sum_{l=0}^{N-1} N \sigma_w^2 |h_l^u|^2 P_l, \quad (59)$$

$$\sum_{k,l,m} \mathbb{E}\left[w_l u_{(l-k)_N}^* w_m^* u_{(m-k)_N}\right] = \sum_{l=0}^{N-1} N \sigma_w^2 |h_l^u|^2 P_l, \quad (60)$$

---

[10] For brevity we define $a_l \triangleq h_l \mu_l$.

[11] Note the term $a_l a_m^* a_{(l-k)_N}^* a_{(m-k)_N}$ is the same if we replace the first with fourth and the second with the third element.

[12] As in (33) and (32), here in (57) in each term, the second and the third symbols are conjugate with the indices $l, (l-k)_N, m, (m-k)_N$ for the first, second, third and the forth term, respectively. For brevity of representations, we have removed the indices as well as the conjugate.

$$\sum_{k,l,m} \mathbb{E}\left[\boldsymbol{u}_l \boldsymbol{w}^*_{(l-k)_N} \boldsymbol{u}^*_m \boldsymbol{w}_{(m-k)_N}\right] = \sum_{l=0}^{N-1} N\sigma_w^2 |h_l^u|^2 P_l, \quad (61)$$

$$\sum_{k,l,m} \mathbb{E}\left[\boldsymbol{u}_l \boldsymbol{u}^*_{(l-k)_N} \boldsymbol{w}^*_m \boldsymbol{w}_{(m-k)_N}\right] = \sum_{l=0}^{N-1} N\sigma_w^2 |h_l^u|^2 P_l, \quad (62)$$

$$\sum_{k,l,m} \mathbb{E}\left[\boldsymbol{u}_l \boldsymbol{u}^*_{(l-k)_N} \boldsymbol{u}^*_m \boldsymbol{u}_{(m-k)_N}\right]$$

$$= 4\sum_{l=0}^{N-1} \delta_{N-1}^l \sum_{m=l+1}^{N-1} |h_l^u|^2 |h_m^u|^2 P_l P_m + \sum_{l=0}^{N-1} |h_l^u|^4 Q_l$$

$$+ 2\sum_{l=0}^{N-1}\sum_{k=1}^{N-1} \Re\{h_l^{u2} \bar{P}_l h_{(l+k)_N}^{u*} h_{(l-k)_N}^{u*} \mu_{(l+k)_N}^* \mu_{(l-k)_N}^*\}$$

$$+ 4\sum_{l=0}^{N-2}\sum_{k=1}^{\frac{N-1}{2}}\sum_{\substack{m=l+1 \\ m\neq(l+k)_N \\ m\neq(l-k)_N}}^{N-1} \Re\{h_l^u h_m^{u*} h_{(l-k)_N}^{u*} h_{(m-k)_N}^u$$

$$\cdot \mu_l \mu_m^* \mu_{(l-k)_N}^* \mu_{(m-k)_N}\} \quad (63)$$

The result of the lemma is obtained by substituting the terms (28) and the baseband equivalent of (29) in (6) altogether.

B. KKT CONDITIONS FOR THE LAGRANGIAN DUAL FUNCTION OF PROBLEM (13)

By writing the Lagrangian for the problem in (13) we have

$$L(\lambda_1, \lambda_2, \zeta_{r0}, \zeta_{i0}, \cdots, \zeta_{r(N-1)}, \zeta_{i(N-1)})$$

$$= \sum_{l=0}^{N-1} -c_0\{\log(1+a_l\sigma_{lr}^2) + \log(1+a_l\sigma_{li}^2)\}$$

$$+ \lambda_1\left(\sum_{l=0}^{N-1} P_l - P^a\right) + \lambda_2\left(P_d - \sum_{l=0}^{N-1} f_{ib}(P_l, \bar{P}_l, \mu_l, h_l, N)\right)$$

$$- \sum_{l=0}^{N-1} \zeta_{lr}\sigma_{lr}^2 + \zeta_{li}\sigma_{li}^2, \quad (64)$$

where $\lambda_1, \lambda_2 \geq 0$ and $\zeta_{lr}, \zeta_{li} \geq 0$ for $l = 0, \cdots, N-1$ are Lagrange multipliers. Accordingly, the KKT conditions read as

$$\sum_{l=0}^{N-1} P_l \leq P_a, \quad (65a)$$

$$\sum_{l=0}^{N-1} f_{ib}(P_l, \bar{P}_l, \mu_l, h_l, N) \geq P_d, \quad (65b)$$

$$\sigma_{lr}^2, \sigma_{li}^2 \geq 0, \ l = 0, ..., N-1, \quad (65c)$$

$$\lambda_1\left(\sum_{l=0}^{N-1} P_l - P_a\right) = 0, \quad (65d)$$

$$\lambda_2\left(P_d - \sum_{l=0}^{N-1} f_{ib}(P_l, \bar{P}_l, \mu_l, h_l, N)\right) = 0, \quad (65e)$$

$$\zeta_{lr}\sigma_{lr}^2 = \zeta_{li}\sigma_{li}^2 = 0, \ l = 0, ..., N-1, \quad (65f)$$

$$\begin{cases} \frac{-c_1 a_l}{1+a_l\sigma_{lr}^2} + \lambda_1 - \lambda_2 f_{ib}^{P_{lr}} - \zeta_{lr} = 0, \ l = 0, ..., N-1 \\ \frac{-c_1 a_l}{1+a_l\sigma_{li}^2} + \lambda_1 - \lambda_2 f_{ib}^{P_{li}} - \zeta_{li} = 0, \ l = 0, ..., N-1 \\ \frac{2c_1 a_l \mu_{lr}}{1+a_l\sigma_{lr}^2} - \lambda_2 f_{ib}^{\mu_{lr}} + 2\zeta_{lr}\mu_{lr} = 0, \ l = 0, ..., N-1 \\ \frac{2c_1 a_l \mu_{li}}{1+a_l\sigma_{li}^2} - \lambda_2 f_{ib}^{\mu_{li}} + 2\zeta_{li}\mu_{li} = 0, \ l = 0, ..., N-1 \end{cases}, \quad (65g)$$

where $c_1 = c_0 \log e$, and

$$\begin{cases} f_{ib}^{P_{lr}} = \alpha_l Q_l^{P_{lr}} + \beta_l + g_1(P_l) \\ \qquad + \Re\left\{\sum_{k=1}^{\frac{N-1}{2}} \mu_{(l+k)_N}^* \mu_{(l-k)_N}^* \Phi_{l,k}\right\} \\ f_{ib}^{P_{li}} = \alpha_l Q_l^{P_{li}} + \beta_l + g_1(P_l) \\ \qquad - \Re\left\{\sum_{k=1}^{\frac{N-1}{2}} \mu_{(l+k)_N}^* \mu_{(l-k)_N}^* \Phi_{l,k}\right\} \\ f_{ib}^{\mu_{lr}} = \alpha_l Q_l^{\mu_{lr}} + \Re\{j2\mu_{li}\sum_{k=1}^{\frac{N-1}{2}} \mu_{(l+k)_N}^* \mu_{(l-k)_N}^* \Phi_{l,k}\} \\ \qquad + \sum_{\substack{d=0 \\ d\neq l}}^{N-1} \Re\{\bar{P}_d T_{l,d}\} \\ \qquad + \sum_{d=0}^{N-2}\sum_{k=1}^{\frac{N-1}{2}}\sum_{\substack{m=d+1 \\ m\neq(d+k)_N \\ m\neq(d-k)_N}}^{N-1} \Re\{\Psi_{d,m,k} S_{d,m,k}^{\mu_{lr}}\} \\ f_{ib}^{\mu_{li}} = \alpha_l Q_l^{\mu_{li}} + \Re\{j2\mu_{lr}\sum_{k=1}^{\frac{N-1}{2}} \mu_{(l+k)_N}^* \mu_{(l-k)_N}^* \Phi_{l,k}\} \\ \qquad - \sum_{\substack{d=0 \\ d\neq l}}^{N-1} \Re\{j\bar{P}_d T_{l,d}\} \\ \qquad + \sum_{d=0}^{N-2}\sum_{k=1}^{\frac{N-1}{2}}\sum_{\substack{m=d+1 \\ m\neq(d+k)_N \\ m\neq(d-k)_N}}^{N-1} \Re\{\Psi_{d,m,k} S_{d,m,k}^{\mu_{li}}\} \end{cases}, \quad (66)$$

where (we removed the details due to the lack of space)

$$g_1(P_l) = \frac{dg(P_l)}{dP_l} = \sum_{\substack{m=0 \\ m\neq l}}^{N-1} \gamma_{m,l} P_m, \quad (67)$$

$$T_{d,l} \triangleq \begin{cases} \mu_{(2d-l)_N}^* \Phi_{l,(d-l)_N} & \text{if } 1 \leq (d-l)_N \leq \frac{N-1}{2} \\ \mu_{(2d-l)_N}^* \Phi_{l,(l-d)_N} & \text{if } 1 \leq (l-d)_N \leq \frac{N-1}{2} \end{cases}, \quad (68)$$

$$Q_l^{P_{lr}} = 6P_{lr} + 2P_{li}, \quad (69)$$

$$Q_l^{P_{li}} = 6P_{li} + 2P_{lr}, \quad (70)$$

$$Q_l^{\mu_{lr}} = -8\mu_{lr}^3, \quad (71)$$

$$Q_l^{\mu_{li}} = -8\mu_{li}^3, \quad (72)$$

and $S_{d,m,k}^{\mu_{lr}}$ and $S_{d,m,k}^{\mu_{li}}$ are defined as below

$$S_{d,m,k}^{\mu_{lr}} = \delta_{l-d} \cdot \mu_m^* \mu_{(l-k)_N}^* \mu_{(m-k)_N}$$

$$+ \delta_{l-m} \cdot \mu_d \mu_{(d-k)_N}^* \mu_{(l-k)_N}$$

$$+ \delta_{l-(l-k)_N} \cdot \mu_d \mu_m^* \mu_{(m-k)_N}$$

$$+ \delta_{l-(m-k)_N} \cdot \mu_d \mu_m^* \mu_{(d-k)_N}, \quad (73)$$

$$S_{d,m,k}^{\mu_{li}} = \delta_{l-d} \cdot j\mu_m^* \mu_{(l-k)_N}^* \mu_{(m-k)_N}$$

$$- \delta_{l-m} \cdot j\mu_d \mu_{(d-k)_N}^* \mu_{(l-k)_N}$$

$$- \delta_{l-(l-k)_N} \cdot j\mu_d \mu_m^* \mu_{(m-k)_N}$$

$$+ \delta_{l-(m-k)_N} \cdot j\mu_d \mu_m^* \mu_{(d-k)_N}. \quad (74)$$

It is verified that λ_1 can not be zero. Because otherwise the first two sets of equations in (65g) for the KKT conditions contradict the equalities. Therefore, the average power constraint is satisfied with equality due to (65d). For $\lambda_2 = 0$, it can be verified that we have the water-filling solution.

C. Proof of Lemma 1

By noting that $\mu_l = 0$, $l = 0,\ldots,N-1$ and writing the Lagrangian KKT conditions in (65), we have

$$\sum_{l=0}^{N-1} P_l \leq P_a, \tag{75a}$$

$$\sum_{l=0}^{N-1} f_{ib}(P_l,\bar{P}_l,0,h_l,N) \geq P_d, \tag{75b}$$

$$P_{lr},P_{li} \geq 0, \; l=0,...,N-1, \tag{75c}$$

$$\lambda_1 \left(\sum_{l=0}^{N-1} P_l - P_a \right) = 0, \tag{75d}$$

$$\lambda_2 \left(P_d - \sum_{l=0}^{N-1} f_{ib}(P_l,\bar{P}_l,0,h_l,N) \right) = 0, \tag{75e}$$

$$\zeta_{lr}P_{lr} = \zeta_{li}P_{li} = 0, \; l=0,...,N-1 \tag{75f}$$

$$\begin{cases} \frac{-c_l a_l}{1+a_l P_{lr}} + \lambda_1 - \lambda_2 f_{ib}^{P_{lr}} - \zeta_{lr} = 0, \; l=0,...,N-1, \\ \frac{-c_l a_l}{1+a_l P_{li}} + \lambda_1 - \lambda_2 f_{ib}^{P_{li}} - \zeta_{li} = 0, \; l=0,...,N-1, \end{cases} \tag{75g}$$

where

$$f_{ib}^{P_{lr}} = \alpha_l Q_l^{P_{lr}} + \beta_l + g_1(P_l) = \alpha_l(6P_{lr}+2P_{li}) + \beta_l + g_1(P_l),$$
$$f_{ib}^{P_{li}} = \alpha_l Q_l^{P_{li}} + \beta_l + g_1(P_l) = \alpha_l(6P_{li}+2P_{lr}) + \beta_l + g_1(P_l).$$

with $g_1(P_l)$, $Q_l^{P_{lr}}$ and $Q_l^{P_{li}}$ defined in (67), (69) and (70), respectively.

It is verified that λ_1 cannot be zero. Because otherwise the equations in (75g) contradict the equalities. Therefore, average power constraint is satisfied with equality. For $\lambda_2 = 0$ it can be verified that we have water-filling solution. For $\lambda_2 \neq 0$, note that ζ_{lr}, ζ_{li}, $l=0,\ldots,N-1$ act as slack variables, so they can be eliminated. We have

$$\sum_{l=0}^{N-1} P_l = P_a, \tag{76a}$$

$$\sum_{l=0}^{N-1} f_{ib}(P_l,\bar{P}_l,0,h_l,N) = P_d, \tag{76b}$$

$$P_{lr},P_{li} \geq 0, \; l=0,...,N-1, \tag{76c}$$

$$P_{lr}\left(\lambda_1 - G_l(P_r^N,P_i^N)\right) = 0, \; l=0,...,N-1, \tag{76d}$$

$$P_{li}\left(\lambda_1 - G_l(P_i^N,P_r^N)\right) = 0, \; l=0,...,N-1, \tag{76e}$$

where due to ζ_{lr}, $\zeta_{li} \geq 0$ for any l, for the optimal solution, we have

$$\lambda_1 \geq \max_l \{G_l(P_r^N,P_i^N), G_l(P_i^N,P_r^N)\}. \tag{77}$$

References

[1] B. Clerckx, R. Zhang, R. Schober, D. W. K. Ng, D. I. Kim, and H. V. Poor, "Fundamentals of wireless information and power transfer: from RF energy harvester models to signal and system designs," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 1, pp. 4–33, Jan. 2019.

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

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

[4] 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.

[5] 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.

[6] 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.

[7] A. S. Boaventura and N. B. Carvalho, "Maximizing DC power in energy harvesting circuits using multisine excitation," *2011 IEEE MTT-S Int'l Microwave Sym.*, pp. 1–4, Jun. 2011.

[8] B. Clerckx, "Wireless information and power transfer: Nonlinearity, waveform design, and Rate-Energy tradeoff," *IEEE Trans. Signal Processing*, vol. 66, no. 4, pp. 847–862, Feb. 2018.

[9] E. Boshkovska, D. W. K. Ng, N. Zlatanov, A. Koelpin, and R. Schober, "Robust resource allocation for MIMO wireless powered communication networks based on a non-linear EH model," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 1984–1999, May 2017.

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

[11] ——, "Low-complexity adaptive multisine waveform design for wireless power transfer," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 2207–2210, 2017.

[12] J. Kim, B. Clerckx, and P. D. Mitcheson, "Signal and system design for wireless power transfer : Prototype, experiment and validation," *CoRR*, vol. abs/1901.01156, 2019. [Online]. Available: http://arxiv.org/abs/1901.01156

[13] B. Clerckx and J. Kim, "On the beneficial roles of fading and transmit diversity in wireless power transfer with nonlinear energy harvesting," *IEEE Trans. on Wireless Commun.*, vol. 17, no. 11, pp. 7731–7743, Nov. 2018.

[14] Y. Zeng, B. Clerckx, and R. Zhang, "Communications and signals design for wireless power transmission," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2264–2290, May 2017.

[15] P. N. Alevizos and A. Bletsas, "Sensitive and nonlinear far-field RF energy harvesting in wireless communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3670–3685, Jun. 2018.

[16] M. Varasteh, B. Rassouli, and B. Clerckx, "Wireless information and power transfer over an AWGN channel: Nonlinearity and asymmetric gaussian signaling," in *IEEE Inf. Theory Workshop (ITW)*, Nov. 2017, pp. 181–185.

[17] ——, "On capacity-achieving distributions over complex AWGN channels under nonlinear power constraints and their applications to SWIPT," *CoRR*, vol. abs/1712.01226, 2017. [Online]. Available: http://arxiv.org/abs/1712.01226

[18] E. Bayguzina and B. Clerckx, "Modulation design for wireless information and power transfer with nonlinear energy harvester modeling," *IEEE 19th Int'l Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, Jun. 2018.

[19] J. Kang, I. Kim, and D. I. Kim, "Wireless information and power transfer: Rate-Energy tradeoff for nonlinear energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1966–1981, Mar. 2018.

[20] K. Xiong, B. Wang, and K. J. R. Liu, "Rate-Energy region of SWIPT for MIMO broadcasting under nonlinear energy harvesting model," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5147–5161, Aug. 2017.

[21] X. Xu, A. Özçelikkale, T. McKelvey, and M. Viberg, "Simultaneous information and power transfer under a non-linear RF energy harvesting model," in *IEEE Int'l Conference on Commun. Workshops (ICC Workshops)*, May 2017, pp. 179–184.

[22] R. Morsi, V. Jamali, D. W. K. Ng, and R. Schober, "On the capacity of SWIPT systems with a nonlinear energy harvesting circuit," *IEEE Int'l Conf. Commun. (ICC)*, pp. 1–7, May 2018.

[23] N. Khalfet and S. M. Perlaza, "Simultaneous information and energy transmission in the two-user gaussian interference channel," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 1, pp. 156–170, Jan. 2019.

[24] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

Morteza Varasteh received the M.S. degree in communications systems engineering from Sharif University of Technology, Tehran, Iran, in 2011. He received his Ph.D. degree in communications from Imperial College London, London, UK in 2016. He is currently a research associate in the Communications and Signal Processing group at Imperial College London. His research interests are in the general area of information theory, wireless communications, optimization theory and machine learning.

Borzoo Rassouli Borzoo Rassouli received the M.Sc. degree in electrical engineering from university of Tehran, Iran in 2012, and the Ph.D. degree in communications engineering from Imperial College London, UK in 2016. He was a postdoctoral research associate at Imperial College from 2016 to 2018. In August 2018, he joined university of Essex as a lecturer (Assistant Professor). His research interests lie in the general areas of information theory and statistics.

Bruno Clerckx [SM'17] received the M.S. and Ph.D. degrees in applied science from the UniversitÃ© Catholique de Louvain, Louvain-la-Neuve, Belgium, in 2000 and 2005, respectively. From 2006 to 2011, he was with Samsung Electronics, Suwon, South Korea, where he actively contributed to 3GPP LTE/LTE-A and IEEE 802.16m and acted as the Rapporteur for the 3GPP Coordinated Multi-Point (CoMP) Study Item. From 2014 to 2016, he was an Associate Professor with Korea University, Seoul, South Korea. He also held visiting research appointments at Stanford University, EURECOM, National University of Singapore, The University of Hong Kong and Princeton University. Since 2011, he has been with Imperial College London, first as a Lecturer from 2011 to 2015, then as a Senior Lecturer from 2015 to 2017, and now as a Reader. He is currently a Reader (Associate Professor) with the Electrical and Electronic Engineering Department, Imperial College London, London, U.K.

He has authored two books, 160 peer-reviewed international research papers, and 150 standards contributions, and is the inventor of 75 issued or pending patents among which 15 have been adopted in the specifications of 4G (3GPP LTE/LTE-A and IEEE 802.16m) standards. His research area is communication theory and signal processing for wireless networks. He has been a TPC member, a symposium chair, or a TPC chair of many symposia on communication theory, signal processing for communication and wireless communication for several leading international IEEE conferences. He is an Elected Member of the IEEE Signal Processing Society SPCOM Technical Committee. He served as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2011 to 2015 and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2014 to 2018, and is currently an Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He has also been a (lead) guest editor for special issues of the EURASIP Journal on Wireless Communications and Networking, IEEE ACCESS and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He was an Editor for the 3GPP LTE-Advanced Standard Technical Report on CoMP.