Energy Efficiency Maximization of Full-Duplex Two-Way Relay with Non-ideal Power Amplifiers and Non-negligible Circuit Power

Qimei Cui, *Senior Member, IEEE*, Yuhao Zhang, Wei Ni, *Senior Member, IEEE*,
Mikko Valkama, *Senior Member, IEEE*, and Riku Jäntti, *Senior Member, IEEE*

Abstract—In this paper, we maximize the energy efficiency (EE) of full-duplex (FD) two-way relay (TWR) systems under non-ideal power amplifiers (PAs) and non-negligible transmission-dependent circuit power. We start with the case where only the relay operates full duplex and two timeslots are required for TWR. Then, we extend to the advanced case, where the relay and the two nodes all operate full duplex, and accomplish TWR in a single timeslot. In both cases, we establish the intrinsic connections between the optimal transmit powers and durations, based on which the original non-convex EE maximization can be convexified and optimally solved. Simulations show the superiority of FD-TWR in terms of EE, especially when traffic demand is high. The simulations also reveal that the maximum EE of FD-TWR is more sensitive to the PA efficiency, than it is to self-cancellation. The full FD design of FD-TWR is susceptible to traffic imbalance, while the design with only the relay operating in the FD mode exhibits strong tolerance.

Index Terms—Energy efficiency, full-duplex, non-ideal power amplifier, non-negligible circuit power, two-way relay.

I. INTRODUCTION

Full duplex (FD) has recently been extensively studied due to significant advance on self-cancellation techniques [1]–[6]. The self-interference from the transmit antenna(s) of a node to the receive antenna(s) of the same node can be suppressed by means of passive suppression (PS) such as antenna separation [1]–[3], and active cancellation (AC) such as radio-frequency circuit designs and digital signal processing [2]–[5]. By collectively using these techniques, a self-cancellation of 80 dB, comprising 40 dB from PS and 40 dB from AC, has been recorded [2], [3].

Two-way relay (TWR) has also drawn extensive recent interest, where advanced digital signal processing has been developed to minimize the number of timeslots required for two mutually hidden nodes to exchange data via a shared

half-duplex (HD) relay [7], [8]. Physical-layer network coding (PNC) has been proposed to achieve the minimum number of two slots under HD settings [8]. TWR also has the advantage of improving energy efficiency (EE), as extensively studied in [9]–[12].

There is an opportunity to join the advantages of FD and TWR, referred to as FD-TWR, to improve the spectral efficiency (SE) and EE of TWR. Practical conditions need to be taken into account, such as limited power supply in mobile nodes, non-ideal power amplifier (PA) and non-negligible circuit power consumption. To this end, the maximization of EE is of practical interest, especially when the relay and hidden nodes are mobile nodes and do not have persistent power supply. Despite the EE maximization has been extensively studied in HD relay-assisted wireless networks [9]–[12], few works have been carried out on HD-TWR under non-ideal PA and non-negligible circuit power, not to mention FD-TWR. To the best of our knowledge, the collective impact of non-ideal PAs, non-negligible circuit power, and self-interference on FD-TWR has not been studied heretofore.

In most cases, the non-ideal PAs and non-negligible circuit power are non-linear to the transmit power. Under prevailing PA models, such as traditional PA (TPA) [13], the PA power is non-linear to the transmit power. In contrast, the non-negligible circuit power is non-linear and non-convex, since it is typically linear to the data rate which takes the logarithm of the transmit power. The non-ideal PAs and non-negligible circuit power necessitate new optimization variables, i.e., transmit durations, apart from the transmit powers. Particularly, if the nodes transmit excessively long, the circuit energy consumption on reception would increase. If the nodes transmit too short, the transmit powers would become excessively high, while the standby energy of the circuits would grow. The EE would degrade in either of these cases.

In this paper, the EE of FD-TWR is maximized under non-ideal PAs and non-negligible rate-dependent circuit power. We start with the case where only the relay is FD enabled and can transmit to one of two end nodes while receiving from the other. Two timeslots are required for TWR (the case is referred to as FD-TWR-2TS). The EE maximization is formulated with intractable non-convex constraints resulting from self-interference at the FD relay. The necessary conditions of the optimal solution are derived to convert the original non-convex problem to a convex problem with respect to (w.r.t) the transmit durations. We further extend the necessary conditions to the advanced FD-TWR design, referred to as FD-TWR-

1TS, where the relay and the two end nodes all operate in the FD mode, and exploit PNC to accomplish TWR within a single timeslot. The EE maximization of FD-TWR-1TS is also proved to be non-convex. The necessary conditions facilitate convexifying this problem while preserving optimality. Corroborated by extensive simulations, the convex reformulation of both FD-TWR-2TS and FD-TWR-1TS can be readily implemented using standard CVX tools.

The contributions of this paper are summarized as follows.

- The necessary conditions of the most energy efficient schedule for FD-TWR, which are key to convexify the problem and preserve optimality;
- Convexification of the EE maximization for FD-TWR under non-ideal PAs and non-negligible circuit powers, which enables efficient solvers to solve the problem with guaranteed convergence and optimality;
- Important findings: (a) the EE of FD-TWR is more sensitive to the PA efficiency, than it is to self-cancellation; (b) FD-TWR-1TS is susceptible to traffic imbalance, while FD-TWR-2TS exhibits strong tolerance; (c) FD-TWR is superior to HD-TWR in terms of EE in the case of high data rate demand.

The rest of the paper is organized as follows. Section II surveys the related works. Section III presents the system model. The EE maximization problems of FD-TWR-2TS and FD-TWR-1TS are formulated and solved in Sections IV and V, respectively. Simulation results are provided in Section VI, followed by a conclusion in Section VII.

II. RELATED WORK

Earlier works [9]–[12] maximized the EE of HD-TWR under the assumption of ideal PA model and negligible circuit power. PAs are typically non-ideal with the PA efficiency changing with the transmit power, and the circuit powers (i.e., baseband processing, radio-frequency (RF) generation and the circuit standby power etc.) are non-negligible, ranging from tens to hundreds of milliWatts [14], [15]. In [16], under the assumption of ideal PA, non-negligible circuit power consumption was considered when the EE of amplify-and-forward (AF) HD-TWR was maximized. The EE was quasi-concave over the transmit power, and maximized by using the Dinkelbachs method. In [17], the aggregated utility of EE and proportional fairness were maximized for OFDMA-based HD-TWR where the circuit power in active sub-channels was evaluated for energy consumption. In [18], the EE maximization of two timeslots decode-and-forward (DF) HD-TWR was formulated under the non-ideal TPA model and solved using a standard CVX toolbox. In [15], the EE maximization was generalized under a variety of non-ideal PA models and HD-TWR strategies (e.g., two and three timeslots TWR). The generalized problems were converted to problems with convex structures in the vicinities of the optimum, and solved with guaranteed optimality.

Only a handful of works have been on focused the EE maximization of FD-TWR, typically under the assumptions of ideal PAs and/or constant/transmission-independent circuit power. In [19], a non-convex problem was cast to maximize the EE of AF FD-TWR given a SE requirement. A suboptimal solution was developed, comprising two alternating steps.

In [20], four typical power-scaling schemes were proposed to improve the EE of multi-pair AF FD-TWR with a massive antenna array at the relay. With the number of antennas going to infinity, the asymptotic SE and EE were derived. In [21], a new Lattice code with structured binning was designed for FD-TWR, where the relay only quantizes RF combined signals of the end nodes by a constellation lattice and forwards the quantized signals. The end nodes can recover their desired signals, exploiting the meticulously designed binning.

Most existing works on the EE of FD have been on one-way relaying, typically under the assumption of ideal PAs and/or constant circuit power consumption. In [22], the closed-form expressions for the achievable ergodic rates were derived for dual-hop FD massive MIMO AF systems. In [23], discrete stochastic optimization and the Dinkelbach method were taken in an alternating manner to maximize the EE of a multi-cell OFDMA network with shared FD relays in coverage overlapping areas. In [24], a selective DF protocol was proposed to select FD nodes for cooperative relay, given a required outage threshold. In [2], a comparison study was conducted on the EE of FD and HD dual-hop AF systems. Opportunistic relaying mode selection in coupling with transmission power adaptation was developed to maximize the EE. In [25], HD relay was shown to outperform FD relay in terms of throughput under a DF mode at a cost of the increased relay buffer. However, this conclusion is inapplicable to advanced PNC, such as the Lattice code (as considered in this paper), where signals are only decoded at the destination. The relay just quantizes and forwards the signals. The requirement of the relay buffer is constant, depending on the quantization delay.

In different yet relevant contexts, the impact of non-ideal PAs or non-negligible circuit power was evaluated in other communication systems. In [13], the EE-delay tradeoff of a proportional-fair downlink cellular system was studied under non-ideal PAs. In [26], the power allocation was optimized to maximize a bits-per-Joule EE in the conventional single-hop frequency-selective channels under the assumption of non-negligible constant circuit power. In [27], a string tautening algorithm was proposed to produce the most energy-efficient schedule for delay-limited traffic, first under the assumption of negligible circuit power, and then extended to non-negligible constant circuit powers [28] and energy-harvesting communications [29]. In [30], the rate region of a FD OFDM link was maximized under non-ideal transceivers by modelling the self-interference as an additive error vector magnitude (EVM) noise which was decomposed into the equivalent noises at the two ends of the FD link. The equivalent SINRs of the two link directions were separately formulated and jointly maximized by taking the sub-gradient method. However, the results of these works are not applicable to FD-TWR, due to distinctive system architectures.

III. SYSTEM MODEL

Consider an in-band FD-TWR point-to-point wireless network [31], [32], where there are two end nodes, termed A and B, and a relay node, termed R, indicated by the subscripts "$_a$", "$_b$" and "$_r$", respectively. Node R lies between nodes A and B. All the nodes are equipped with directional antennas. Nodes A and B do not have a direct link, and need to exchange

traffic through node R at the same frequency [19], [20]. This scenario is typical in practice, since the three nodes are unaligned in many cases and the directivity of the directional antennas prevents nodes A and B from having a direct link. In many other cases, nodes A and B are too far away and beyond reach. The three nodes can operate in different modes, depending on the FD-TWR strategies specified shortly. The system bandwidth is W. Flat-fading channels are assumed with complex channel coefficients $h_{i,j}$ $(i,j \in \{a,b,r\})$ from node i to j. Particularly, $h_{i,i}$ is the complex channel coefficient of the self-interference channel at node i, from the transmit antenna of the node to its own receive antenna. The additive white Gaussian noise (AWGN) has zero mean and variance σ_j^2, denoted by $n_j \sim \mathcal{CN}(0, \sigma_j^2)$.

A. Full-Duplex and Self-interference Cancellation

Under the FD mode, the self-interference can be partly suppressed by PS and/or AC. Generally, PS, on its own, can account for over 40 dB reduction of self-interference. Working together with AC, it can suppress up to 80 dB. The joint use of PS and AC is referred to as PSAC [1]. We denote the ratio of the self-interference before and after self-cancellation as α_{PS} and α_{PSAC} under PS and PSAC, respectively.

The residual self-interference, after self-cancellation, is typically modeled as a zero-mean Gaussian random variable [33], [34]. The residual self-interference at node i is $e_i = \sqrt{P_i}\tilde{h}_{i,i}x_i$, where $\tilde{h}_{i,i}$ is the residual interference channel coefficient after self-cancellation at node i, and x_i $(E\{|x_i|^2\} = 1)$ is the transmit signal of the node. The variance of the residual self-interference is $P_i\left|\tilde{h}_{i,i}\right|^2$, where $\left|\tilde{h}_{i,i}\right|^2$ is dependent on the specific cancellation method implemented at node i, e.g., PS or PSAC [35].

B. Practical Energy Consumption

The efficiency of a PA is defined to be the ratio between the desired average transmit power and the actual power consumed at the PA. In practice, PAs are non-ideal, and the PA efficiency changes (in many cases, nonlinearly) along with the output transmit power. The representative models of non-ideal PA are TPA and envelope-tracking PA (ETPA). The energy consumptions of TPA and ETPA are given by [13]

$$\Psi_{TPA}(P_i) \approx \frac{\sqrt{P_i P_{MAX,i}}}{\eta_{MAX,i}}, \quad i \in \{a, r, b\}, \qquad (1)$$

$$\Psi_{ETPA}(P_i) \approx \frac{P_i + u P_{MAX,i}}{(1+u)\eta_{MAX,i}}, \quad i \in \{a, r, b\}, \qquad (2)$$

where P_i is the mean output transmit power of node i, $P_{MAX,i}$ and $\eta_{MAX,i}$ are the maximum PA output power and the maximum PA efficiency of node i, respectively, and $u \approx 0.0082$ is a PA dependent parameter for ETPA [13].

Let $P_{\max,i}$ denote the maximum average transmit power of node i, which is important to the calculation of energy consumption. Typically, $P_{\max,i}$ is around $7 \sim 8$ dB lower than $P_{MAX,i}$, since the peak-to-average-power ratio (PAPR), defined by $\kappa_i = P_{MAX,i}/P_{\max,i}$, is around $7 \sim 8$ dB in modern communication systems. Also let $\eta_{\max,i}$ denote the

(a) HD-TWR-2TS with PNC.

(b) FD-TWR-2TS with FD transmission.
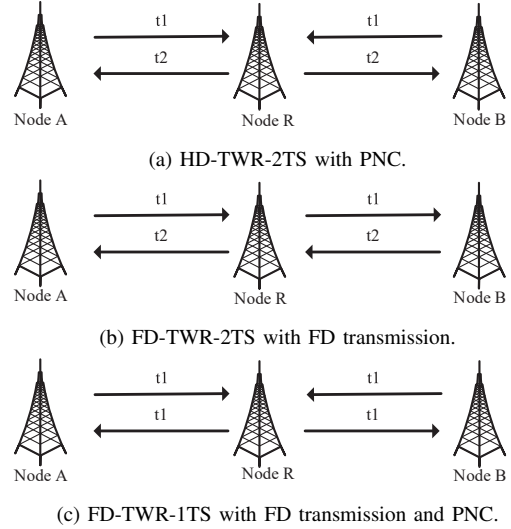

(c) FD-TWR-1TS with FD transmission and PNC.

Fig. 1. Illustration of the three transmission strategies.

PA efficiency associated with $P_{\max,i}$. We can rewrite (1) and (2), w.r.t. $P_{\max,i}$, as

$$\Psi_{TPA}(P_i) \approx \frac{\sqrt{P_i P_{\max,i}}}{\eta_{\max,i}}, \quad i \in \{a, r, b\}, \qquad (3)$$

$$\Psi_{ETPA}(P_i) \approx \frac{P_i + u \kappa_i P_{\max,i}}{(1 + u \kappa_i)\eta_{\max,i}}, \quad i \in \{a, r, b\}, \qquad (4)$$

where, in the case of TPA, $\eta_{\max,i} = \eta_{MAX,i}/\sqrt{\kappa_i}$ is derived by setting the right-hand sides (RHSs) of (1) and (3) equal and solving the equation for $\eta_{\max,i}$; in the case of ETPA, $\eta_{\max,i} = \frac{1+u}{1+u\kappa_i}\eta_{MAX,i}$ is derived by setting the RHSs of (2) and (4) equal and solving the equation for $\eta_{\max,i}$. Clearly, the ideal PA is a special case of ETPA by letting $u = 0$.

The practical circuit power is also non-negligible and transmission-dependent. It can be decomposed into a static component $P_{base,i}$, which drives hardware; and a dynamic component $P_{c,i}$, which accounts for analog and digital signal processing. $P_{c,i} = \varepsilon R$ depends on the transmit (receive) data rate R, where the coefficient ε is the power consumption per unit data rate [36]. As such, the total transmit circuit power can be written as

$$P_{tx,i} = \Psi[P_i(R)] + \varepsilon R + P_{base,i}, \quad i \in \{a, r, b\}. \qquad (5)$$

The total receive circuit power can be written as

$$P_{rx,i} = \varepsilon R + P_{base,i}, \quad i \in \{a, r, b\}. \qquad (6)$$

The power consumption of a node being in the idle mode (neither transmitting or receiving) is set to be constant, denoted by $P_{idle,i}$, due to its independence of P_i.

C. TWR Transmission Strategies

Two strategies of FD-TWR are considered, namely, FD-TWR-1TS and FD-TWR-2TS, as shown in Fig. 1. FD-TWR-2TS is relatively basic, where only the node R is enabled with FD, and nodes A and B operate in the HD mode. As a result, two timeslots are required to complete TWR. FD-TWR-1TS is much advanced, where not only node R, but nodes A and B are also FD enabled. As a result, only a single timeslot is required to accomplish TWR.

Consider a fixed time frame of T seconds, within which a round of TWR is accomplished. Let t_k $(k \in \{1\}$ in FD-TWR-1TS, and $k \in \{1, 2\}$ for FD-TWR-2TS) define the duration of the k-th timeslot within T, as shown in Fig. 1, where $\sum_{\forall k} t_k \leq T$. The nodes switch to an idle mode after $\sum_{\forall k} t_k$ till the end of the frame.

For ease of understanding, we start with FD-TWR-2TS, then move to FD-TWR-1TS.

1) FD-TWR-2TS: In this case, two timeslots are required. Nodes A and B, each equipped with a single antenna, transmit and receive in an alternating manner in the two timeslots, using the HD mode. Node R, equipped with two antennas (one for transmission and one for reception), transmits and receives simultaneously in each timeslot using FD. Specifically, in the first of the two timeslots, node A transmits data (x_a) to the node R which receives and immediately forwards the data (x_r^a) to node B; and the other way around $(x_b$ and $x_r^b)$ in the second timeslot. The DF mode is used in FD-TWR-2TS. There can be a one-frame delay from node R receiving the signals till node R forwarding the signals, but this pipeline delay does not compromise the system capacity.

With the self-interference (partly) cancelled, the signal received at node R in the first timeslot can be given by

$$y_r = \sqrt{P_a} h_{a,r} x_a + \sqrt{P_{r,b}} \tilde{h}_{r,r} x_r^a + n_r. \tag{7}$$

The signal received at node B in the first timeslot can be given by

$$y_b = \sqrt{P_{r,b}} h_{r,b} x_r^a + n_b. \tag{8}$$

Likewise, the signals received at nodes R and A in the second timeslot can be given respectively by

$$y_r = \sqrt{P_b} h_{b,r} x_b + \sqrt{P_{r,a}} \tilde{h}_{r,r} x_r^b + n_r, \tag{9}$$

$$y_a = \sqrt{P_{r,a}} h_{r,a} x_r^b + n_a, \tag{10}$$

where $\tilde{h}_{r,r} = \sqrt{1/\alpha} h_{r,r}$, $E\{|x_a|^2\} = 1$, $E\{|x_b|^2\} = 1$, $E\{|x_r^a|^2\} = 1$, and $E\{|x_r^b|^2\} = 1$.

As discussed in Section III-A, the residual self-interference at node R, $e_r^i = \sqrt{P_{r,j}} \tilde{h}_{r,r} x_r^i$, yields $\mathcal{CN}(0, P_{r,j}|\tilde{h}_{r,r}|^2)$, where $i \in \{a, b\}$ and $j \in \{b, a\}$. This is because the transmit and receive antennas of a FD node are typically isolated, or carefully calibrated to eliminate the line-of-sight (LOS). Moreover, AC can whiten self-interference during signal processing [33], [34]. As a result, the channel between the transmitter and the receiver at a node typically follows the Rayleigh fading [5]. The channel capacities $C_{a,r}$ and $C_{r,b}$ in the forward link direction in the first timeslot satisfy

$$C_{a,r} = \frac{t_1}{T} W \log_2(1 + \frac{P_a |h_{a,r}|^2}{P_{r,b}|\tilde{h}_{r,r}|^2 + \sigma_r^2}), \tag{11}$$

$$C_{r,b} = \frac{t_1}{T} W \log_2(1 + \frac{P_{r,b}|h_{r,b}|^2}{\sigma_b^2}). \tag{12}$$

Likewise, the capacities $C_{b,r}$ and $C_{r,a}$ in the reverse link in the second timeslot satisfy

$$C_{b,r} = \frac{t_2}{T} W \log_2(1 + \frac{P_b |h_{b,r}|^2}{P_{r,a}|\tilde{h}_{r,r}|^2 + \sigma_r^2}), \tag{13}$$

$$C_{r,a} = \frac{t_2}{T} W \log_2(1 + \frac{P_{r,a}|h_{r,a}|^2}{\sigma_a^2}). \tag{14}$$

Finally, the actual achievable data rate in each of the two link directions can be given by

$$C_{a,b} = \min\{C_{a,r}, C_{r,b}\}, \ C_{b,a} = \min\{C_{b,r}, C_{r,a}\}. \tag{15}$$

2) FD-TWR-1TS: In this case, only a single timeslot is required, as mentioned earlier. The three nodes, each equipped with a transmit antenna and a receive antenna, operate in the FD mode during the timeslot. Nodes A and B transmit x_a and x_b to node R, respectively, meanwhile node R transmits x_r to the nodes A and B. Additionally, PNC is adopted to generate x_r as a function of x_a and x_b [10], [18], [34].

The received signals at the three nodes can be given by

$$y_r = \sqrt{P_a} h_{a,r} x_a + \sqrt{P_b} h_{b,r} x_b + \sqrt{P_r} \tilde{h}_{r,r} x_r + n_r, \tag{16}$$

$$y_a = \sqrt{P_r} h_{r,a} x_r + \sqrt{P_a} \tilde{h}_{a,a} x_a + n_a, \tag{17}$$

$$y_b = \sqrt{P_r} h_{r,b} x_r + \sqrt{P_b} \tilde{h}_{b,b} x_b + n_b, \tag{18}$$

where $E\{x_i\} = 0$; $E\{|x_i|^2\} = 1$ $(i \in \{a, b, r\})$; $\sqrt{P_i} \tilde{h}_{i,i} x_i$ is the residual self-interference at node $i \in \{a, b, r\}$; and $e_i = \sqrt{P_i} \tilde{h}_{i,i} x_i$ yields $\mathcal{CN}(0, P_i|\tilde{h}_{i,i}|^2)$.

We consider nested lattice code and structured binning, as defined by three n-dimensional lattices Λ_C^n, Λ_1^n and Λ_2^n. Assume $\Lambda_1^n \subseteq \Lambda_2^n \subseteq \Lambda_C^n$. Λ_C^n is used as codewords, and the Voronoi regions of Λ_1^n and Λ_2^n are the shaping regions of nodes A and B. Node R also uses Λ_1^n as its shaping region. Rather than recovering the exact signals of nodes A and B, node R recovers a structured binned version of the signals combined, by evaluating its Euclidean distance to the codewords [21], exploiting maximum-likelihood estimation [37], [38], or minimum mean square error (MMSE) [39]. On the receipt of the binned signals, nodes A and B can recover the desired signal from each other by canceling the instantaneous self-interference and subtracting their own signals from the binned signals [21].

There can be a one-frame delay from node R receiving the signals till it forwarding the binned version of the signals. In other words, the instantaneous self-interference is independent of the signals that nodes A and B subtract from their received binned signals, as described in [21]. Such a delay is a pipeline delay, and does not compromise the throughput or EE of the network under stable channel conditions. This paper focuses on the allocation of temporal and energy resources to maximize the EE in FD-TWR systems using nested lattice coding. Designing nested lattice codes is beyond the scope of the paper.

Considering the nested lattice coding with structured binning [21], the channel capacities from nodes A and B to node R can be given by

$$C_{a,r} = \frac{t_1}{T} W \log_2\left(\frac{P_a |h_{a,r}|^2}{P_a |h_{a,r}|^2 + P_b |h_{b,r}|^2} + \frac{P_a |h_{a,r}|^2}{P_r |\tilde{h}_{r,r}|^2 + \sigma_r^2}\right), \tag{19}$$

$$C_{b,r} = \frac{t_1}{T} W \log_2\left(\frac{P_b |h_{b,r}|^2}{P_a |h_{a,r}|^2 + P_b |h_{b,r}|^2} + \frac{P_b |h_{b,r}|^2}{P_r |\tilde{h}_{r,r}|^2 + \sigma_r^2}\right). \tag{20}$$

The data rates from node R to nodes A and B can be respectively given by

$$C_{r,a} = \frac{t_1}{T} W \log_2(1 + \frac{P_r|h_{r,a}|^2}{P_a\left|\tilde{h}_{a,a}\right|^2 + \sigma_a^2}), \qquad (21)$$

$$C_{r,b} = \frac{t_1}{T} W \log_2(1 + \frac{P_r|h_{r,b}|^2}{P_b\left|\tilde{h}_{b,b}\right|^2 + \sigma_b^2}), \qquad (22)$$

where $\left(P_i\left|\tilde{h}_{i,i}\right|^2 + \sigma_i^2\right)$ is the variance of $(e_i + n_i)$, due to the fact that e_i and n_i are independent.

Substituting (19)-(22) into (15), we can finally obtain the achievable data rate in both link directions.

IV. EE MAXIMIZATION OF FD-TWR-2TS

In this section, we maximize the EE of FD-TWR-2TS, given the data rate requirements R_{rl} and R_{fl}, under non-ideal PAs and non-negligible circuit power. Here, R_{rl} and R_{fl} are the data rate requirements in the reverse link (i.e., from node B to node A) and the forward link (i.e., from node A to node B), respectively.

The EE is defined as the ratio between the average total data rate in both link directions and the average total power consumed by the nodes [26], as given by

$$\eta_E = \frac{R_{rl} + R_{fl}}{E_{\text{total}}/T} = \frac{(R_{rl} + R_{fl})T}{E_{\text{total}}}, \qquad (23)$$

where the second equality indicates that the EE is equivalent to the ratio between the number of bits to be transmitted (in both directions) within T and the total energy required to transmit the bits, denoted by E_{total}. To this end, given the number of bits to be sent within T, i.e., $(R_{rl} + R_{fl})T$, maximizing η_E is equivalent to minimizing E_{total}.

Minimizing E_{total} facilitates maximizing the EE, under non-negligible circuit power. The reason is that, under non-negligible circuit power, maximizing the EE may require the nodes to transmit for part of a frame to leverage the non-negligible signal processing energy and circuit standby energy. The optimal transmit rate may switch to null during a frame. The direct maximization of the EE, i.e., directly maximizing η_E, would be unsuitable, due to such change of the data rate. Therefore, the EE maximization can be formulated as

$$\min_{P_a, P_b, P_{r,a}, P_{r,b}, t_1, t_2} E_{\text{total}} \qquad \textbf{(P1)}$$

$$\text{s.t.} \quad \min\{C_{b,r}, C_{r,a}\} \geq R_{rl}, \ \min\{C_{a,r}, C_{r,b}\} \geq R_{fl};$$

$$0 \leq \Psi(P_i) \leq P_{\max,i}, \ i \in \{a, b, r\};$$

$$t_1 + t_2 \leq T, \ 0 \leq t_1 \leq T, \ 0 \leq t_2 \leq T,$$

where the total energy consumption $E_{\text{total}} = (P_{tx,a} + P_{tx,r} + P_{rx,b} + P_{rx,r})t_1 + (P_{tx,b} + P_{tx,r} + P_{rx,a} + P_{rx,r})t_2 + P_{idle}(T - t_1 - t_2)$ and $P_{idle} = P_{idle,a} + P_{idle,b} + P_{idle,r}$.

(P1) is non-convex w.r.t. the transmit powers and durations, i.e., P_i ($i \in \{a, b, r\}$) and t_j ($j \in \{1, 2\}$). This is because that under both TPA and ETPA, the logarithmic data constraint is non-convex in P_i and t_j, due to the non-convexity nature of logarithm. This leads to a non-convex feasible set for the EE maximization problem of interest. Further, under TPA, the objective of **(P1)** is non-convex, as can be rigorously verified through the Hessian matrix **H**, as given by

$$\mathbf{H} = \begin{bmatrix} -\Gamma_1(a)t_1 & 0 & 0 & \Gamma_2(a) & 0 \\ 0 & -\Gamma_1(r)(t_1+t_2) & 0 & \Gamma_2(r) & \Gamma_2(r) \\ 0 & 0 & -\Gamma_1(b)t_2 & 0 & \Gamma_2(b) \\ \Gamma_2(a) & \Gamma_2(r) & 0 & 0 & 0 \\ 0 & \Gamma_2(r) & \Gamma_2(b) & 0 & 0 \end{bmatrix}$$

where

$$\Gamma_1(i) = \frac{P_{\max,i}^2}{4\eta_{\max,i}(P_i P_{\max,i})^{2/3}} > 0, \ i \in \{a, r, b\},$$

$$\Gamma_2(i) = \frac{P_{\max,i}}{2\eta_{\max,i}\sqrt{P_i P_{\max,i}}} > 0, \ i \in \{a, r, b\}.$$

Clearly, **H** is not positive definite, given negative leading principal minors, e.g., the first and third leading principal minors.

To convexify **(P1)**, we can prove that the necessary condition of the optimal solution for **(P1)** can be given by

$$C_{b,r} = C_{r,a} = R_{rl}, \ C_{a,r} = C_{r,b} = R_{fl}. \qquad (24)$$

The detailed proof is provided in Appendix A.

As a result, the optimal transmit powers, P_a^*, P_b^*, $P_{r,b}^*$ and $P_{r,a}^*$, can be rewritten as the functions of the optimal transmit durations, t_1^* and t_2^*, as given by

$$P_{r,b}^*(t_1^*) = \frac{\sigma_b^2}{|h_{r,b}|^2}\left(2^{\frac{R_{fl}T}{t_1^*W}} - 1\right); \qquad (25)$$

$$P_{r,a}^*(t_2^*) = \frac{\sigma_a^2}{|h_{r,a}|^2}\left(2^{\frac{R_{rl}T}{t_2^*W}} - 1\right); \qquad (26)$$

$$P_a^*(t_1^*) = \frac{\sigma_r^2}{|h_{a,r}|^2}\left(2^{\frac{R_{fl}T}{t_1^*W}} - 1\right) + \frac{\sigma_b^2\left|\tilde{h}_{r,r}\right|^2}{|h_{a,r}|^2|h_{r,b}|^2}\left(2^{\frac{R_{fl}T}{t_1^*W}} - 1\right)^2; \qquad (27)$$

$$P_b^*(t_2^*) = \frac{\sigma_r^2}{|h_{b,r}|^2}\left(2^{\frac{R_{rl}T}{t_2^*W}} - 1\right) + \frac{\sigma_a^2\left|\tilde{h}_{r,r}\right|^2}{|h_{b,r}|^2|h_{r,a}|^2}\left(2^{\frac{R_{rl}T}{t_2^*W}} - 1\right)^2. \qquad (28)$$

Therefore, **(P1)** can be reformulated to only optimize the transmit durations, t_i ($i = 1, 2$), as given by

$$\min_{\mathbf{t}} \quad E_{\text{total}} = E(\mathbf{t}) \qquad (29)$$

$$\text{s.t.} \quad \mathbf{t}_{\min} \leq \mathbf{t};$$

$$\mathbf{1}^T \mathbf{t} \leq T,$$

where $E(\cdot)$ is the objective function denoting the overall energy consumption. $\mathbf{t}$ is a column vector collecting both the time durations to be optimized, $\mathbf{t}_{min}$ is a column vector collecting the minimum value of each time duration $t_{\min,i}$, and $\mathbf{1}$ is the all-one column vector. The minimum time durations are specified by $t_i \geq t_{\min,i}$, which can be readily calculated.

In the case of FD-TWR-2TS, $\mathbf{t}_{\min} = [t_{\min,1}, t_{\min,2}]^T$. $t_{\min,1}$ and $t_{\min,2}$ are obtained by separately setting (25)–(28) to be no larger than the corresponding maximum transmit powers, solving the inequalities, and choosing the intersection of the results.

For notation simplicity, we denote $\lambda_{fl} = \frac{R_{fl}T}{Wt_1}$ and $\lambda_{rl} = \frac{R_{rl}T}{Wt_2}$. λ_{fl} and λ_{rl} are the spectrum efficiency in the forward

and reverse link directions, respectively. The following theorems dictate that (29) yields a convex structure under non-ideal PAs and non-negligible circuit power.

Theorem 1: *Under TPA, (29) is convex for FD-TWR-2TS.*

Proof: The proof starts by substituting (25)–(28) into the objective function of (**P1**). The objective function can be rewritten as

$$
E\left(t_1, t_2\right) =
$$

$$
\left\{ \sqrt{\beta_{1,1} \times 2^{\frac{R_{fl}T}{t_1 W}} + \gamma_1 \times \left(2^{\frac{R_{fl}T}{t_1 W}}\right)^2} + \sqrt{\alpha_{1,1} \times 2^{\frac{R_{fl}T}{t_1 W}}} \right\} t_1 +
$$

$$
\left\{ \sqrt{\beta_{1,2} \times 2^{\frac{R_{rl}T}{t_2 W}} + \gamma_2 \times \left(2^{\frac{R_{rl}T}{t_2 W}}\right)^2} + \sqrt{\alpha_{1,2} \times 2^{\frac{R_{rl}T}{t_2 W}}} \right\} t_2 +
$$

$$
P_{1,1} t_1 + P_{1,2} t_2 + (P_{idle,a} + P_{idle,b} + P_{idle,r}) T,
$$

$$
\alpha_{1,1} = \frac{P_{\max,r}\sigma_b^2}{\eta_{\max,r}^2 |h_{r,b}|^2}, \quad \alpha_{1,2} = \frac{P_{\max,r}\sigma_a^2}{\eta_{\max,r}^2 |h_{r,a}|^2},
$$

$$
\beta_{1,1} = \frac{P_{\max,a}\sigma_r^2}{\eta_{\max,a}^2 |h_{a,r}|^2}, \quad \beta_{1,2} = \frac{P_{\max,b}\sigma_r^2}{\eta_{\max,b}^2 |h_{b,r}|^2},
$$

$$
\gamma_1 = \frac{P_{\max,a}\sigma_b^2 \left|\tilde{h}_{r,r}\right|^2}{\eta_{\max,a}^2 |h_{a,r}|^2 |h_{r,b}|^2}, \quad \gamma_2 = \frac{P_{\max,b}\sigma_a^2 \left|\tilde{h}_{r,r}\right|^2}{\eta_{\max,b}^2 |h_{b,r}|^2 |h_{r,a}|^2},
$$

where

$$
P_{1,1} = P_{base,a} + P_{base,b} + 2P_{base,r} + 4\varepsilon R_{fl} - P_{idle,a} - P_{idle,b} - P_{idle,r},
$$

$$
P_{1,2} = P_{base,a} + P_{base,b} + 2P_{base,r} + 4\varepsilon R_{rl} - P_{idle,a} - P_{idle,b} - P_{idle,r}.
$$

$E(t_1, t_2)$ is convex, because Hessian matrix is positive definite, as given by

$$
\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}
$$

$$
A = \frac{1}{4\alpha_{1,1}^2 q_1^{\frac{3}{2}} t_1^3 W^2} [\beta_{1,1}^2 \sqrt{\alpha_{1,1} 2^{\lambda_{fl}}} + \gamma_1 2^{\lambda_{fl}} (4p_1 + \alpha_{1,1}\sqrt{q_1})
$$
$$
+ \beta_{1,1}(6p_1 + \alpha_{1,1}\sqrt{q_1})](\alpha_{1,1} 2^{\lambda_{fl}})^{\frac{3}{2}} R_{fl}^2 T^2 \log(2)^2 > 0,
$$

$$
B = \frac{1}{4\alpha_{1,2}^2 q_2^{\frac{3}{2}} t_2^3 W^2} [\beta_{1,2}^2 \sqrt{\alpha_{1,2} 2^{\lambda_{rl}}} + \gamma_2 2^{\lambda_{rl}} (4p_2 + \alpha_{1,2}\sqrt{q_2})
$$
$$
+ \beta_{1,2}(6p_2 + \alpha_{1,2}\sqrt{q_2})](\alpha_{1,2} 2^{\lambda_{rl}})^{\frac{3}{2}} R_{rl}^2 T^2 \log(2)^2 > 0,
$$

where $p_1 = \frac{(\alpha_{1,1} 2^{\lambda_{fl}})^{\frac{3}{2}} \gamma_1}{\alpha_{1,1}}$, $q_1 = 2^{\lambda_{fl}} \left(\beta_{1,1} + \gamma_1 2^{\lambda_{fl}}\right)$, $p_2 = \frac{(\alpha_{1,2} 2^{\lambda_{rl}})^{\frac{3}{2}} \gamma_2}{\alpha_{1,2}}$, $q_2 = 2^{\lambda_{rl}} \left(\beta_{1,2} + \gamma_2 2^{\lambda_{rl}}\right)$ ∎

Theorem 2: *Under ETPA, (29) is convex for FD-TWR-2TS.*

Proof: To prove this, we substitute (25)–(28) into the objective function of (**P1**). The objective function can be written as

$$
E\left(t_1, t_2\right) = \left\{ \alpha_{2,1} \left(2^{\frac{R_{fl}T}{t_1 W}} - 1\right) + \beta_{2,1} \left(2^{\frac{R_{fl}T}{t_1 W}} - 1\right)^2 \right\} t_1
$$

$$
+ \left\{ \alpha_{2,2} \left(2^{\frac{R_{rl}T}{t_2 W}} - 1\right) + \beta_{2,2} \left(2^{\frac{R_{rl}T}{t_2 W}} - 1\right)^2 \right\} t_2
$$

$$
+ P_{2,1} t_1 + P_{2,2} t_2 + (P_{idle,a} + P_{idle,b} + P_{idle,r}) T,
$$

$$
\alpha_{2,1} = \frac{\sigma_r^2}{(1 + u\kappa_a)\eta_{\max,a}|h_{a,r}|^2} + \frac{\sigma_b^2}{(1 + u\kappa_r)\eta_{\max,r}|h_{r,b}|^2},
$$

$$
\alpha_{2,2} = \frac{\sigma_r^2}{(1 + u\kappa_b)\eta_{\max,b}|h_{b,r}|^2} + \frac{\sigma_a^2}{(1 + u\kappa_r)\eta_{\max,r}|h_{r,a}|^2},
$$

$$
\beta_{2,1} = \frac{\sigma_b^2 \left|\tilde{h}_{r,r}\right|^2}{(1 + u\kappa_a)\eta_{\max,a}|h_{a,r}|^2 |h_{r,b}|^2},
$$

$$
\beta_{2,2} = \frac{\sigma_a^2 \left|\tilde{h}_{r,r}\right|^2}{(1 + u\kappa_b)\eta_{\max,b}|h_{b,r}|^2 |h_{r,a}|^2},
$$

where

$$
P_{2,1} = \frac{u\kappa_a P_{\max,a}}{(1 + u\kappa_a)\eta_{\max,a}} + \frac{u\kappa_r P_{\max,r}}{(1 + u\kappa_r)\eta_{\max,r}} + P_{base,a} + P_{base,b}
$$
$$
+ 2P_{base,r} + 4\varepsilon R_{fl} - P_{idle,a} - P_{idle,b} - P_{idle,r},
$$

$$
P_{2,2} = \frac{u\kappa_b P_{\max,b}}{(1 + u\kappa_b)\eta_{\max,b}} + \frac{u\kappa_r P_{\max,r}}{(1 + u\kappa_r)\eta_{\max,r}} + P_{base,a} + P_{base,b}
$$
$$
+ 2P_{base,r} + 4\varepsilon R_{rl} - P_{idle,a} - P_{idle,b} - P_{idle,r}.
$$

Similarly, the Hessian matrix of $E\left(t_1, t_2\right)$ can be proved to be positive definite. As a result, $E\left(t_1, t_2\right)$ is convex. ∎

V. EE MAXIMIZATION OF FD-TWR-1TS

In this section, we consider FD-TWR-1TS under non-ideal PAs and non-negligible circuit powers. The EE maximization of FD-TWR-1TS can be formulated as

$$
\min_{P_a, P_b, P_r, t_1} \quad E_{\text{total}} \tag{P2}
$$

$$
\text{s.t.} \quad \min\{C_{b,r}, C_{r,a}\} \geq R_{rl}, \quad \min\{C_{a,r}, C_{r,b}\} \geq R_{fl};
$$

$$
0 \leq \Psi(P_i) \leq P_{\max,i}, \quad i \in \{a, b, r\};
$$

$$
0 \leq t_1 \leq T,
$$

where the total energy consumption is $E_{\text{total}} = (P_{tx,a} + P_{tx,b} + P_{tx,r} + P_{rx,a} + P_{rx,b} + P_{rx,r})t_1 + P_{idle}(T - t_1)$ and $P_{idle} = P_{idle,a} + P_{idle,b} + P_{idle,r}$.

(**P2**) is not convex due to the non-convex feasible region posed by non-convex logarithmic data rate constraints, as well as the non-linear transmit power constraints as discussed for FD-TWR-2TS in Section III. As proved in Appendix B, the necessary conditions of the optimal solution for (**P2**) can be given by

$$
C_{b,r} = R_{rl}, C_{a,r} = R_{fl}; \tag{30}
$$

$$
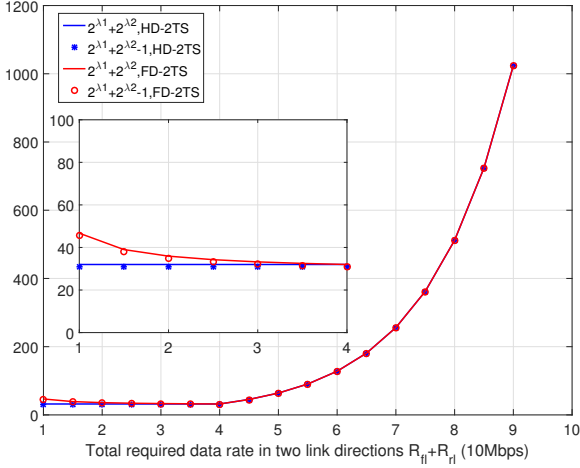\min\{C_{r,a} - R_{rl}, C_{r,b} - R_{fl}\} = 0. \tag{31}
$$

Fig. 2. Example for the approximate relationship justification where the simulation parameters can be found in Table I ($\lambda 1 = \lambda_{fl}$ and $\lambda 2 = \lambda_{rl}$).

We are particularly interested in the asymptotic case where the data rate requirements are high, i.e., $2^{\lambda_{fl}} + 2^{\lambda_{rl}} \gg 1$. This consideration is reasonable, since FD is typically exploited to increase data rates. In the case where the data rate requirements are not high, this approximation of $2^{\lambda_{fl}} + 2^{\lambda_{rl}} \gg 1$ can also hold to a great extent. This is because the transmit time of the nodes, i.e., t_1, is reduced to decrease circuit energy consumption and increase the EE. As a result, $\frac{T}{t_1}$ increases with the decrease of $\frac{R_{fl}}{W}$ and $\frac{R_{rl}}{W}$, avoiding significant decreases of $\lambda_{fl} = \frac{R_{fl}}{W}\frac{T}{t_1}$ and $\lambda_{rl} = \frac{R_{rl}}{W}\frac{T}{t_1}$. This approximation is reliable over a wide range of R_{fl} and R_{rl}, as shown in Fig. 2.

In this case, solving (19), (20) and (30), we can write the optimal transmit powers, P_a and P_b, as the functions of t_1 and P_r, as given by

$$P_a(t_1) = 2^{\lambda_{fl}} \frac{(2^{\lambda_{fl}} + 2^{\lambda_{rl}} - 1)(P_r(t_1)\left|\tilde{h}_{r,r}\right|^2 + \sigma_r^2)}{(2^{\lambda_{fl}} + 2^{\lambda_{rl}})|h_{a,r}|^2}$$
$$\approx 2^{\lambda_{fl}} \frac{P_r(t_1)\left|\tilde{h}_{r,r}\right|^2 + \sigma_r^2}{|h_{a,r}|^2}, \quad (32)$$

$$P_b(t_1) = 2^{\lambda_{rl}} \frac{(2^{\lambda_{fl}} + 2^{\lambda_{rl}} - 1)(P_r(t_1)\left|\tilde{h}_{r,r}\right|^2 + \sigma_r^2)}{(2^{\lambda_{fl}} + 2^{\lambda_{rl}})|h_{b,r}|^2}$$
$$\approx 2^{\lambda_{rl}} \frac{P_r(t_1)\left|\tilde{h}_{r,r}\right|^2 + \sigma_r^2}{|h_{b,r}|^2}. \quad (33)$$

From (31), it can be seen that P_r is the larger of $P_{r,rl}$ and $P_{r,fl}$, which is the solution for the equality of (31) with (21) and (22), respectively. Therefore,

$$P_r(t_1) = \max\{P_{r,rl}; P_{r,fl}\}. \quad (34)$$

We proceed with two cases to derive the explicit transmit powers of all the three nodes, i.e., nodes R, A, and B. First, $P_r(t_1)$ is derived, and then substituted to (32) and (33) for $P_a(t_1)$ and $P_b(t_1)$.

CASE I: In the case of $P_{r,rl} \geq P_{r,fl}$, (21) meets the equality for the optimal solution. P_r is solved through the equalities of (21) and (32), as given by

$$P_{r,1}(t_1) = P_{r,rl} \approx \frac{2^{\lambda_{rl}}\left(2^{\lambda_{fl}}\left|\tilde{h}_{a,a}\right|^2\sigma_r^2 + |h_{a,r}|^2\sigma_a^2\right)}{|h_{r,a}|^2|h_{a,r}|^2 - 2^{\lambda_{fl}}2^{\lambda_{rl}}\left|\tilde{h}_{a,a}\right|^2\left|\tilde{h}_{r,r}\right|^2}. \quad (35)$$

Note that the transmit power $P_{r,1}(t_1) \geq 0$ is positive in practical, thus $|h_{r,a}|^2|h_{a,r}|^2 - 2^{\lambda_{fl}}2^{\lambda_{rl}}\left|\tilde{h}_{a,a}\right|^2\left|\tilde{h}_{r,r}\right|^2 > 0$. Given the channel data requirements, this imposes minimum requirement of the self-cancellation level.

Substitute (35) into (32) and (33). We then have

$$P_{a,1}(t_1) \approx \frac{2^{\lambda_{fl}}\left(2^{\lambda_{rl}}\left|\tilde{h}_{r,r}\right|^2\sigma_a^2 + |h_{r,a}|^2\sigma_r^2\right)}{|h_{r,a}|^2|h_{a,r}|^2 - 2^{\lambda_{fl}}2^{\lambda_{rl}}\left|\tilde{h}_{a,a}\right|^2\left|\tilde{h}_{r,r}\right|^2}, \quad (36)$$

$$P_{b,1}(t_1) \approx \frac{2^{\lambda_{rl}}|h_{a,r}|^2\left(2^{\lambda_{rl}}\left|\tilde{h}_{r,r}\right|^2\sigma_a^2 + |h_{r,a}|^2\sigma_r^2\right)}{\left(|h_{r,a}|^2|h_{a,r}|^2 - 2^{\lambda_{fl}}2^{\lambda_{rl}}\left|\tilde{h}_{a,a}\right|^2\left|\tilde{h}_{r,r}\right|^2\right)|h_{b,r}|^2}. \quad (37)$$

CASE II: In the case of $P_{r,rl} < P_{r,fl}$, (22) meets the equality for the optimal solution. P_r is solved through the equalities of (22) and (33), as given by

$$P_{r,2}(t_1) = P_{r,fl} \approx \frac{2^{\lambda_{fl}}\left(2^{\lambda_{rl}}\left|\tilde{h}_{b,b}\right|^2\sigma_r^2 + |h_{b,r}|^2\sigma_b^2\right)}{|h_{r,b}|^2|h_{b,r}|^2 - 2^{\lambda_{fl}}2^{\lambda_{rl}}\left|\tilde{h}_{b,b}\right|^2\left|\tilde{h}_{r,r}\right|^2}. \quad (38)$$

Substitute (38) into (32) and (33). We then have

$$P_{a,2}(t_1) \approx \frac{2^{\lambda_{fl}}|h_{b,r}|^2\left(2^{\lambda_{fl}}\left|\tilde{h}_{r,r}\right|^2\sigma_b^2 + |h_{r,b}|^2\sigma_r^2\right)}{\left(|h_{r,b}|^2|h_{b,r}|^2 - 2^{\lambda_{fl}}2^{\lambda_{rl}}\left|\tilde{h}_{b,b}\right|^2\left|\tilde{h}_{r,r}\right|^2\right)|h_{a,r}|^2}, \quad (39)$$

$$P_{b,2}(t_1) \approx \frac{2^{\lambda_{rl}}\left(2^{\lambda_{fl}}\left|\tilde{h}_{r,r}\right|^2\sigma_b^2 + |h_{r,b}|^2\sigma_r^2\right)}{|h_{r,b}|^2|h_{b,r}|^2 - 2^{\lambda_{fl}}2^{\lambda_{rl}}\left|\tilde{h}_{b,b}\right|^2\left|\tilde{h}_{r,r}\right|^2}. \quad (40)$$

Now, we can obtain the equivalent optimization problem of (**P2**) in the same form as (29), with the objective changed to

$$E(t_1) = \max\{E_1(t_1), E_2(t_1)\}, \quad (41)$$

where $E_1(t)$ and $E_2(t)$ are the energy consumption in *CASE I* and *CASE II*, respectively.

In the case of FD-TWR-1TS, $\mathbf{t}_{\min} = t_{\min,1}$. $t_{\min,1}$ can be obtained by setting (35)–(40) to be no larger than the corresponding maximum transmit powers, solving the inequalities, and choosing the intersection of the results. Specifically, we can rearrange the inequalities to be polynomial with all positive coefficients. Given the monotonicity of such polynomials, one-dimensional search, such as the bisection method, can be used to efficiently achieve $t_{\min,1}$.

Theorem 3: *Under TPA, (29) is convex for FD-TWR-1TS in the case of high data rate demands.*

Proof: In the asymptotic case that $2^{\lambda_{fl}} + 2^{\lambda_{rl}} \gg 1$, exploiting the TPA model, $E_i(t_1)$ can be written as

$$E_i(t_1) = \sqrt{\frac{P_{\max,a}}{\eta_{\max,a}^2}} t_1 \sqrt{P_{a,i}(t_1)} + \sqrt{\frac{P_{\max,b}}{\eta_{\max,b}^2}} t_1 \sqrt{P_{b,i}(t_1)}$$
$$+ \sqrt{\frac{P_{\max,r}}{\eta_{\max,r}^2}} t_1 \sqrt{P_{r,i}(t_1)} + P_3 t_1 + P_{idle} T,$$

where $i \in \{1,2\}$ and $P_3 = P_{rx} + P_{base} + \varepsilon(R_{fl} + 2R_{rl}) - P_{idle}$.

$E_i(t_1)$ is the positively weighted sum of $t_1\sqrt{P_{a,i}(t_1)}$, $t_1\sqrt{P_{b,i}(t_1)}$ and $t_1\sqrt{P_{r,i}(t_1)}$. We proceed to prove that each of these three components is convex.

CASE I: In the case of $P_{r,rl} \geq P_{r,fl}$, $t_1\sqrt{P_{a,1}(t_1)}$ and $t_1\sqrt{P_{r,1}(t_1)}$ can be confirmed to be strictly convex, since

$$\frac{\partial^2\left(t_1\sqrt{P_{a,1}(t_1)}\right)}{\partial t_1^2} > 0 \,, \quad \frac{\partial^2\left(t_1\sqrt{P_{r,1}(t_1)}\right)}{\partial t_1^2} > 0.$$

Next, we prove the convexity of $t_1\sqrt{P_{b,1}(t_1)}$. To do this, we can rewrite $t_1\sqrt{P_{b,1}(t_1)}$ as

$$t_1\sqrt{P_{b,1}(t_1)} =$$
$$t_1 \sqrt{\frac{2^{\frac{R_{rl}T}{Wt_1}}|h_{a,r}|^2\left(2^{\frac{R_{rl}T}{Wt_1}}\left|\tilde{h}_{r,r}\right|^2\sigma_a^2+|h_{r,a}|^2\sigma_r^2\right)}{\left(|h_{r,a}|^2|h_{a,r}|^2-2^{\frac{(R_{rl}+R_{fl})T}{Wt_1}}\left|\tilde{h}_{a,a}\right|^2\left|\tilde{h}_{r,r}\right|^2\right)|h_{b,r}|^2}}$$
$$= d_3 \times x\sqrt{\frac{2^{\frac{1}{x}}\left(1+a_3 2^{\frac{1}{x}}\right)}{1-b_3 2^{\frac{c_3}{x}}}} = d_3 \times f_{TPA}(x),$$

where $x = (Wt_1)/(R_{rl}T) > 0$, and for notational simplicity,

$$a_3 = \frac{\left|\tilde{h}_{r,r}\right|^2\sigma_a^2}{|h_{r,a}|^2\sigma_r^2} > 0, \quad b_3 = \frac{\left|\tilde{h}_{a,a}\right|^2\left|\tilde{h}_{r,r}\right|^2}{|h_{r,a}|^2|h_{a,r}|^2} > 0,$$
$$c_3 = \frac{(R_{rl}+R_{fl})}{R_{rl}} > 1, \quad d_3 = \sqrt{\frac{\sigma_r^2}{|h_{b,r}|^2}}\frac{R_{rl}T}{W} > 0.$$

To prove the convexity of $t_1\sqrt{P_{b,1}(t_1)}$ now becomes to prove $f_{TPA}(x)$ is convex over $x \in (0, \infty)$. The second-order derivative of $f_{TPA}(x)$ can be given by

$$\frac{\partial^2 f_{TPA}(x)}{\partial x^2} = 2^{-2+\frac{1}{x}}\log(2)^2 \times \qquad (42)$$
$$\frac{\left[g_3(c_3)+4^{\frac{1}{x}}a_3 g_1(c_3)+2^{1+\frac{1}{x}}a_3 g_2(c_3)\right]}{\left(1+2^{\frac{1}{x}}a_3\right)\sqrt{\frac{2^{\frac{1}{x}}\left(1+a_3 2^{\frac{1}{x}}\right)}{1-b_3 2^{\frac{c_3}{x}}}}\left(1-b_3 2^{\frac{c_3}{x}}\right)^3 x^3},$$

where the coefficients are obtained through mathematic manipulation, as given by

$$g_1(c_3) = 4+4^{\frac{c_3}{x}}b_3^2(c_3-2)^2+2^{\frac{c_3+x}{x}}b_3\left(c_3^2+2c_3-4\right),$$
$$g_2(c_3) = 3+4^{\frac{c_3}{x}}b_3^2\left(c_3^2-3c_3+3\right)+2^{\frac{c_3+x}{x}}b_3\left(2c_3^2+3c_3-6\right),$$
$$g_3(c_3) = 1+4^{\frac{c_3}{x}}b_3^2(c_3-1)^2+2^{\frac{c_3+x}{x}}b_3\left(c_3^2+c_3-1\right).$$

We first set $c_3^2+2c_3-4 > 0$, $c_3^2-3c_3+3 > 0$, $2c_3^2+3c_3-6 > 0$, and $c_3^2+c_3-1 > 0$, and then solve the inequalities. We

see that all the four polynomials are positive for $c_3 \in (\sqrt{5}-1, \infty)$. We prove that the second-order derivative of $f_{TPA}(x)$ is positive over $c_3 \in (\sqrt{5}-1, \infty)$.

We can also meticulously rewrite $g_1(c_3)$ and $g_2(c_3)$ as

$$g_1(c_3) = (c_3-2)^2 \times$$
$$\left(\left(2^{\frac{c_3}{x}}b_3+\frac{c_3^2+4c_3-8}{2(c_3-2)^2}\right)^2-\frac{c_3^2\left(c_3^2+4c_3-8\right)}{(c_3-2)^4}\right),$$

$$g_2(c_3) = \left(c_3^2-3c_3+3\right) \times$$
$$\left(\left(2^{\frac{c_3}{x}}b_3+\frac{2c_3^2+3c_3-6}{2\left(c_3^2-3c_3+3\right)}\right)^2-\frac{c_3^2\left(4c_3^2+12c_3-27\right)}{4(c_3^2-3c_3+3)^2}\right).$$

A sufficient condition for $g_i(c_3)$, $i \in \{1,2,3\}$, all being positive, is that $c_3^2+4c_3-8 < 0$, $4c_3^2+12c_3-27 < 0$, and $c_3^2+c_3-1 > 0$. We can solve the inequalities and obtain the feasible region, $c_3 \in (1, 2\sqrt{3}-2)$, given $c_3 > 1$. In other words, the second-order derivative of $f_{TPA}(x)$ is positive when $c_3 \in (1, 2\sqrt{3}-2)$.

As a result, we prove $\frac{\partial^2 f_{TPA}(x)}{\partial x^2} > 0$ over $c_3 \in (1, 2\sqrt{3}-2)\bigcup(\sqrt{5}-1, \infty) = (1, \infty)$. Therefore, $f_{TPA}(x)$ is convex. In turn, $t_1\sqrt{P_{b,1}(t_1)}$ is convex, and so is $E_1(t_1)$, the positively weighted sum of $t_1\sqrt{P_{a,1}(t_1)}$, $t_1\sqrt{P_{b,1}(t_1)}$ and $t_1\sqrt{P_{r,1}(t_1)}$.

CASE II: In the case of $P_{r,rl} < P_{r,fl}$, $t_1\sqrt{P_{b,2}}$ and $t_1\sqrt{P_{r,2}}$ are convex, as proved in *CASE I*. We rewrite $t_1\sqrt{P_{a,2}}$ as

$$t_1\sqrt{P_{a,2}(t_1)} =$$
$$t_1 \sqrt{\frac{2^{\frac{R_{fl}T}{Wt_1}}|h_{b,r}|^2\left(2^{\frac{R_{fl}T}{Wt_1}}\left|\tilde{h}_{r,r}\right|^2\sigma_b^2+|h_{r,b}|^2\sigma_r^2\right)}{\left(|h_{r,b}|^2|h_{b,r}|^2-2^{\frac{(R_{rl}+R_{fl})T}{Wt_1}}\left|\tilde{h}_{b,b}\right|^2\left|\tilde{h}_{r,r}\right|^2\right)|h_{a,r}|^2}}$$
$$= \sqrt{\frac{\sigma_r^2}{|h_{a,r}|^2}}\frac{R_{fl}T}{W}y\sqrt{\frac{2^{\frac{1}{y}}\left(1+\frac{\left|\tilde{h}_{r,r}\right|^2\sigma_b^2}{|h_{r,b}|^2\sigma_r^2}2^{\frac{1}{y}}\right)}{1-2^{\frac{(R_{rl}+R_{fl})}{R_{fl}y}}\frac{\left|\tilde{h}_{b,b}\right|^2\left|\tilde{h}_{r,r}\right|^2}{|h_{r,b}|^2|h_{b,r}|^2}}},$$

where $y = (Wt_1)/(R_{fl}T) > 0$ for notational simplicity.

We can see that $t_1\sqrt{P_{a,2}(t_1)}$ has the same structure as $t_1\sqrt{P_{b,1}(t_1)}$. The only difference is the positive coefficients. To this end, we can evaluate the convexity of $t_1\sqrt{P_{a,2}(t_1)}$ in the same way, as we did on $t_1\sqrt{P_{b,1}(t_1)}$, and prove the convexity of $t_1\sqrt{P_{a,2}(t_1)}$. The detailed proof is suppressed due to limited space.

Given that both $E_1(t_1)$ and $E_2(t_1)$ are convex, $E(t_1) = \max\{E_1(t_1), E_2(t_1)\}$ is convex under the TPA model. $\blacksquare$

Theorem 4: *Under ETPA, (29) is convex for FD-TWR-1TS in the case of high data rate demands.*

Proof: In the asymptotic case that $2^{\lambda_{fl}} + 2^{\lambda_{rl}} \gg 1$, exploiting the ETPA model, $E_i(t_1)$ can be written as

$$E_i(t_1) = \frac{t_1 P_{a,i}(t_1)}{(1+u\kappa_a)\eta_{\max,a}} + \frac{t_1 P_{b,i}(t_1)}{(1+u\kappa_b)\eta_{\max,b}}$$
$$+ \frac{t_1 P_{r,i}(t_1)}{(1+u\kappa_r)\eta_{\max,r}} + P_4 t_1 + P_{idle} T,$$

where $i \in \{1, 2\}$ and $P_4 = \frac{u\kappa_a P_{\max,a}}{(1+u\kappa_a)\eta_{\max,a}} + \frac{u\kappa_b P_{\max,b}}{(1+u\kappa_b)\eta_{\max,b}} + \frac{u\kappa_r P_{\max,r}}{(1+u\kappa_r)\eta_{\max,r}} + P_{rx} + P_{base} + \varepsilon\left(R_{fl} + 2R_{rl}\right) - P_{idle}$.

$E_i(t_1)$ is the positively weighted sum of $t_1 P_{a,i}(t_1)$, $t_1 P_{b,i}(t_1)$ and $t_1 P_{r,i}(t_1)$. We proceed to prove that each of these three components is convex.

CASE I: In the case of $P_{r,rl} \geq P_{r,fl}$, $t_1 P_{a,1}(t_1)$ and $t_1 P_{r,1}(t_1)$ are both convex, as can be readily extended from the proof of Theorem 3. $t_1 P_{b,1}(t_1)$ can be written as

$$t_1 P_{b,1}(t_1) =$$
$$t_1 \frac{2^{\frac{R_{rl}T}{Wt_1}} |h_{a,r}|^2 \left(2^{\frac{R_{rl}T}{Wt_1}} \left|\tilde{h}_{r,r}\right|^2 \sigma_a^2 + |h_{r,a}|^2 \sigma_r^2\right)}{\left(|h_{r,a}|^2 |h_{a,r}|^2 - 2^{\frac{(R_{rl}+R_{fl})T}{Wt_1}} \left|\tilde{h}_{a,a}\right|^2 \left|\tilde{h}_{r,r}\right|^2\right) |h_{b,r}|^2}$$

$$= e_4 \times x \frac{2^{\frac{1}{x}}\left(1 + a_4 2^{\frac{1}{x}}\right)}{1 - b_4 2^{\frac{c_4}{x}}} = e_4 \times f_{ETPA}(x),$$

where $x = (Wt_1)/(R_{rl}T) > 0$, and for notational simplicity,

$$a_4 = \frac{\left|\tilde{h}_{r,r}\right|^2 \sigma_a^2}{|h_{r,a}|^2 \sigma_r^2} > 0, \quad b_4 = \frac{\left|\tilde{h}_{a,a}\right|^2 \left|\tilde{h}_{r,r}\right|^2}{|h_{r,a}|^2 |h_{a,r}|^2} > 0,$$

$$c_4 = \frac{(R_{rl} + R_{fl})}{R_{rl}} > 1, \quad e_4 = \frac{\sigma_r^2}{|h_{b,r}|^2} \frac{R_{rl}T}{W} > 0.$$

To prove the convexity of $t_1 P_{b,1}(t_1)$ now becomes to prove $f_{ETPA}(x)$ is convex over $x \in (0, \infty)$. The second-order derivative of $f_{ETPA}(x)$ can be expressed as

$$\frac{\partial^2 f_{ETPA}(x)}{\partial x^2} = 2^{\frac{1}{x}} \log(2)^2 \times$$
$$\frac{\left[1 + 4^{\frac{c_4}{x}} b_4^{~2}(c_4-1)^2 + 2^{\frac{c_4}{x}} b_4\left(c_4^{~2} + 2c_4 - 2\right) + 2^{\frac{1}{x}} a_4 g_4(c_4)\right]}{\left(1 - b_4 2^{\frac{c_4}{x}}\right)^3 x^3},$$
$$(43)$$

where $g_4(c_4)$ is given by

$$g_4(c_4) = 4 + 4^{\frac{c_4}{x}} b_4^{~2}(c_4 - 2)^2 + 2^{\frac{c_4}{x}} b_4\left(c_4^{~2} + 4c_4 - 8\right).$$

We first set $c_4^{~2} + 2c_4 - 2 > 0$ and $c_4^{~2} + 4c_4 - 8 > 0$, and solve the inequalities. We obtain $c_4 \in (2\sqrt{3} - 2, \infty)$; in other words, we prove that the second-order derivative of $f_{ETPA}(x)$ is positive over $c_4 \in (2\sqrt{3} - 2, \infty)$.

We can also rewrite $g_4(c_4)$ as

$$g_4(c_4) = (c_4 - 2)^2 \times$$
$$\left(\left(2^{\frac{c_4}{x}} b_4 + \frac{c_4^{~2} + 4c_4 - 8}{2(c_4-2)^2}\right)^2 - \frac{c_4^{~2}\left(c_4^{~2} + 8c_4 - 16\right)}{4(c_4-2)^4}\right).$$

We then set $c_4^{~2} + 2c_4 - 2 > 0$ and $c_4^{~2} + 8c_4 - 16 < 0$. We solve the inequalities and obtain $c_4 \in (1, 4\sqrt{2} - 4)$, given $c_4 > 1$. In other words, the second-order derivative of $f_{ETPA}(x)$ is positive for $c_4 \in (1, 4\sqrt{2} - 4)$.

As a result, $\frac{\partial^2 f_{ETPA}(x)}{\partial x^2} > 0$ over $c_4 \in (1, 4\sqrt{2} - 4) \bigcup (2\sqrt{3} - 2, \infty) = (1, \infty)$. $f_{ETPA}(x)$ is convex. Therefore, $t_1 P_{b,1}(t_1)$ is convex, and so is $E_1(t_1)$, as well.

CASE II: In the case of $P_{r,rl} < P_{r,fl}$, $t_1 P_{b,2}(t_1)$ and $t_1 P_{r,2}(t_1)$ are convex due to the fact that the second-order derivatives of these are positive. $t_1 P_{a,2}(t_1)$ can be rewritten as given by

$$t_1 P_{a,2}(t_1) =$$
$$t_1 \frac{2^{\frac{R_{fl}T}{Wt_1}} |h_{b,r}|^2 \left(2^{\frac{R_{fl}T}{Wt_1}} \left|\tilde{h}_{r,r}\right|^2 \sigma_b^2 + |h_{r,b}|^2 \sigma_r^2\right)}{\left(|h_{r,b}|^2 |h_{b,r}|^2 - 2^{\frac{(R_{rl}+R_{fl})T}{Wt_1}} \left|\tilde{h}_{b,b}\right|^2 \left|\tilde{h}_{r,r}\right|^2\right) |h_{a,r}|^2}$$

$$= \frac{\sigma_r^2}{|h_{a,r}|^2} \frac{R_{fl}T}{W} y \frac{2^{\frac{1}{y}}\left(1 + \frac{\left|\tilde{h}_{r,r}\right|^2 \sigma_b^2}{|h_{r,b}|^2 \sigma_r^2} 2^{\frac{1}{y}}\right)}{1 - 2^{\frac{(R_{rl}+R_{fl})}{R_{fl} y}} \frac{|\tilde{h}_{b,b}|^2 |\tilde{h}_{r,r}|^2}{|h_{r,b}|^2 |h_{b,r}|^2}},$$

where $y = \frac{Wt_1}{R_{fl}T} > 0$.

We see that $t_1 P_{a,2}(t_1)$ has the same structure as $t_1 P_{b,1}(t_1)$. The only difference is the positive coefficients. We can evaluate the convexity of $t_1 P_{a,2}(t_1)$ in the same way, as we have done on $t_1 P_{b,1}(t_1)$, and prove the convexity of $t_1 P_{a,2}(t_1)$. The detailed proof is suppressed due to limited space.

Given that both $E_1(t_1)$ and $E_2(t_1)$ are convex, $E(t_1) = \max\{E_1(t_1), E_2(t_1)\}$ is convex under the ETPA model. $\blacksquare$

VI. SIMULATION RESULTS

In this section, simulations are carried out to evaluate the maximum EE of FD-TWR. As dictated in Theorems 1, 2, 3 and 4, the non-convex EE maximization of FD-TWR under non-ideal PAs and non-negligible circuit power can be reformulated to convex optimizations with optimality preserved and variables reduced (i.e., w.r.t. the transmit durations only). The convex equivalents can be solved using standard solvers, such as the barrier method, the exterior penalty method and the sequential quadratic programming (SQP) method. For illustration purpose, we use the barrier method.

For comparison purpose, we also simulate the maximum EE of HD-TWR (the details are provided in Appendix C). The frame duration is $T = 10$ ms, unless otherwise specified. The PA efficiency $\eta_{\max,i}$ is set to 0.35 [40], which is within the reasonable range between 0.311 and 0.388 [41]. Reciprocal channels are assumed between every pair of nodes, i.e., $h_{i,j} = h_{j,i}$ for all $i \neq j$. The noise is independent but identically distributed (i.i.d.) at all the nodes, i.e., $\sigma_i^2 = \sigma^2$, $\forall i$. The PA consumption simulated is in the order of dozens of Watts [40], depending on the distance of the transmitter and receiver, as well as the efficiency of the PA. The circuit power ranges from tens to hundreds of milliWatts, comprising mainly baseband processing and RF generation [40]. Other simulation parameters are specified in Table I, with reference to other works [15], [18], [30] and 3GPP LTE specifications [42]. For fair comparison, we ensure that all schemes involved achieve the same required data rates in both link directions over the fixed time frame T.

Fig. 3 plots the maximum EE of FD-TWR, as the self-cancellation capability of FD-TWR improves. The HD-TWR is plotted in blue for reference purpose. We can see FD-TWR-1TS can substantially outperform FD-TWR-2TS under a wide spectrum of self-cancellation capability, while the gap between the two schemes decreases as the self-cancellation improves. As shown in the figures, when the self-cancellation level is

TABLE I
SIMULATION PARAMETERS

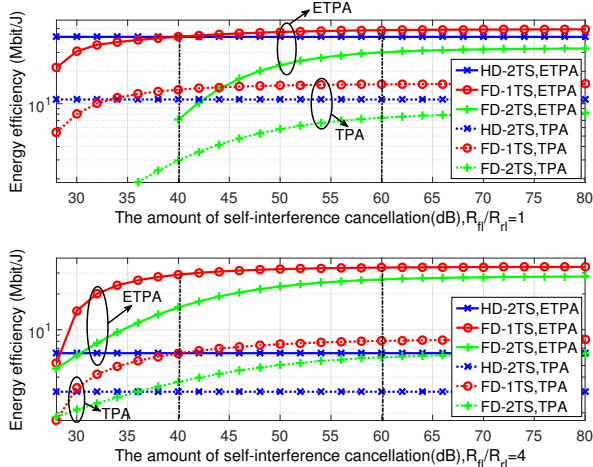| Parameters	Values
System bandwidth W	10 MHz
Noise power spectral density (N_0)	-174 dBm/Hz
Distance between nodes ($d_{a,r}$, $d_{r,b}$)	50, 50 m
Distance between antennas ($d_{a,a}$, $d_{b,b}$, $d_{r,r}$)	5, 5, 5 cm
Average path loss between nodes in dB ($L_{i,j}$)	$103.8 + 21 \log_{10} d$
Cancellation amount (α_{PS}, α_{PSAC})	40, 60 dB
Idle power consumption ($P_{\mathrm{idle},a}$, $P_{\mathrm{idle},r}$, $P_{\mathrm{idle},b}$)	30, 15, 5 mW
Static circuit power consumption ($P_{base,i}$)	100, 50, 20 mW
Dynamic circuit factor (ε)	50 mW/Gbps
Maximum output power ($P_{\max,i}$)	46, 37, 23 dBm
Maximum PA efficiency ($\eta_{\max,i}$)	0.35, 0.35, 0.35

Fig. 3. The optimal EE versus the self-cancellation capability of FD-TWR, where $R_{fl} + R_{rl} = 65$ Mbps.

Fig. 4. EE comparison between different transmission strategies w.r.t data rate demand under the ETPA model.

high, i.e., the self-interference becomes negligible, the EEs of both FD-TWR-1TS and -2TS stabilize and flat out. As the self-cancellation level decreases, FD-TWR-2TS exhibits faster degradation than FD-TWR-1TS. This is because FD-TWR-1TS allows every node to transmit over the entire time frame. The transmit power of the nodes can be lower than that in FD-TWR-2TS, and in turn, produces less self-interference. In this sense, weak self-cancellation capability can still keep the nodes to receive properly in FD-TWR-1TS. In contrast, the transmit durations of the source/destination nodes are much shorter in FD-TWR-2TS. The required transmit powers of the nodes are much higher, due to the exponential-linear tradeoff between the transmit power and time. This results in strong self-interference, which cannot be properly cancelled at the low self-cancellation level. In other words, FD-TWR-2TS is more sensitive to the self-cancellation level.

We notice that HD-TWR can be more energy efficient than FD-TWR, when the relay and the source/destination nodes have poor self-cancellation capability, e.g., ≤ 40 dB if the traffic is balanced in the forward and reverse links, as shown in the upper figure of Fig. 3. The reason is that the residual self-interference is too large and results in a substantial loss of data rate. In turn, the EE of FD-TWR degrades.

We also notice that unbalanced traffic deteriorates the EE degradation, and HD-TWR is much more susceptible to the traffic imbalance than FD-TWR. For FD-TWR-1TS, this is because the transmission and reception of all the FD nodes are aligned within the same time duration. The EE of the un-
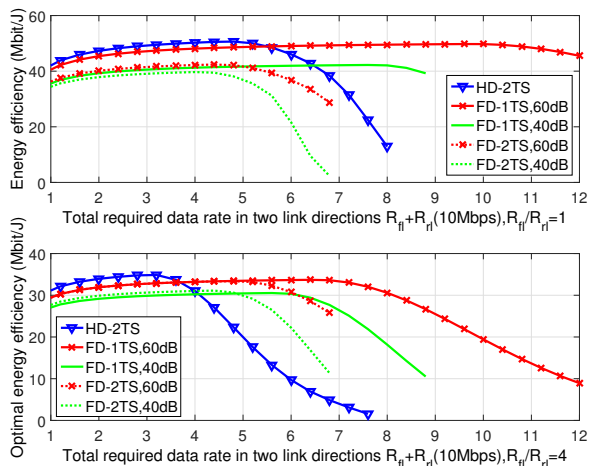
balanced links decreases, because of excessively high transmit power over relatively short time in the high-rate direction, and excessively long transmission time incurring excessive circuit energy consumption in the low rate direction. For HD-TWR, the transmissions of the relay to both source/destination nodes also need to be aligned. Compared to FD-TWR-1TS, HD-TWR has much shorter time durations in which transmissions are aligned, incurring higher transmit power and lower EE, as shown in the lower figure of Fig. 3. Interestingly, FD-TWR-2TS is not much affected by the traffic imbalance. This is because the FD relay receives and relays the same traffic from one node at any instant. The increasingly unbalanced traffic only enlarges difference of time durations allocated to the two link directions at the relay, and does not affect the transmit power of the relay.

Fig. 4 plots the maximum EE of FD-TWR-1TS and -2TS, as well as HD-TWR, as the total rate requirement increases under the self-cancellation levels of 40 and 60 dB. We see that FD-TWR is superior to HD-TWR in terms of EE when the data requirement is high. Particularly, FD-TWR-1TS can maintain the same EE of 50 Mbits/J as HD-TWR, while doubling the data rate region of HD-TWR, from 55 Mbps to 110 Mbps. We also see that the self-cancellation level plays an important role in FD-TWR. Not only does it affect the maximum EE, but also affects the data rate that can be supported by the network. Reducing the level leads to a loss of both the EE and data rate. The traffic imbalance between the two link directions can further reduce the EE and data rate.

An interesting finding in Fig. 4 is that HD-TWR can outperform both the FD-TWR schemes, when the data requirement is low. This is the case where the traffic demand is low and the radio resource is abundant. Only part of the time frame is allocated to data transmission in HD-TWR, and the energy can be most efficiently planned to leverage the PA consumption, circuit consumption, and data transmission. In contrast, FD-TWR has to align the transmission and reception of the FD nodes within the same time duration, incurring the loss of EE, as discussed in Fig. 3.

Fig. 5 corroborates the above findings by evaluating the optimized transmit powers and time durations. We see in
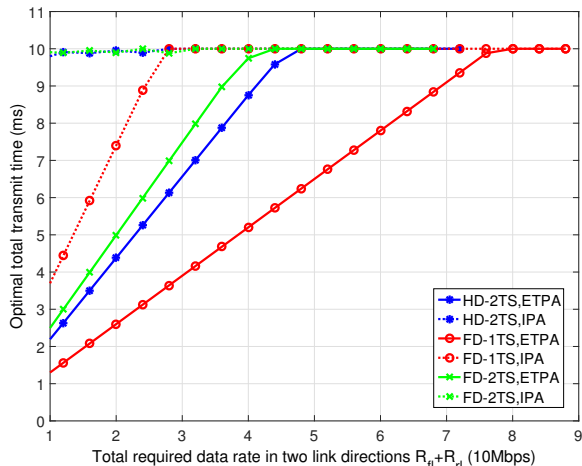
(a) The optimized transmit time under the ETPA model
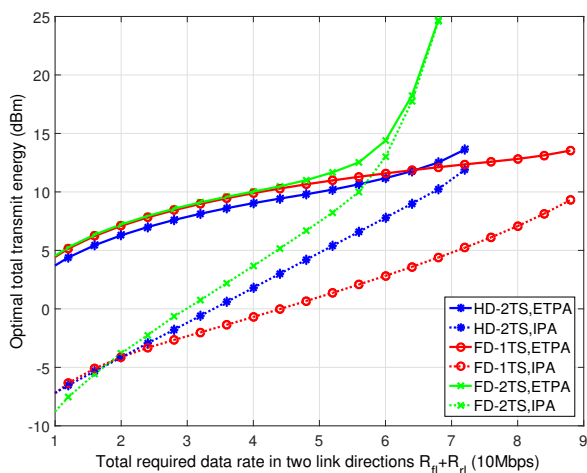
(b) The optimized transmit power under the ETPA model

Fig. 5. The optimal resource allocation results under non-ideal PA and non-negligible circuit power, where PS ($\alpha_{PS} = 40$ dB) is adopted and $R_{fl}/R_{rl} = 1$.

Fig. 5(a) that, when the traffic demand is less than 50 Mbps, the average transmit time within a frame is less than the entire frame for HD-TWR. This is the case where HD-TWR has the chance to separately optimize the transmit power of each link (the transmit power is yet to increase exponentially, as evident from Fig. 5(b)), achieving the higher EE than FD-TWR.

We note that the optimal transmit time under non-negligible transmission-dependent circuit power is less than the entire time frame under low data requirements. The conclusion drawn is that the proposed algorithm can optimize the transmit time as such that the energy consumed for data transmission (including the PA consumption and transmission-dependent circuit operation) is leveraged with the energy consumed in the idle-state circuit.

We also notice that the optimal transmit time of FD-TWR-1TS is not always across the entire frame, even in the ideal case of negligible circuit power and ideal PAs. The reason is that FD-TWR-1TS needs to operate at a relatively high transmit power to overcome both the self-interference and noise. If the transmission lasts over the entire frame, the transmit power would be lower. The self-interference would

also decrease. The receiver noise would become dominant over the self-interference, which could require extra transmit power to suppress the noise given the required data rate and thus reduce the EE.

Fig. 6 shows the impact of PA efficiency, self-cancellation and traffic imbalance on the EE maximization, where the ETPA model is considered for illustration purpose. From Fig. 6(a), we can see that FD-TWR can be much more tolerant to the imbalance of traffic than HD-TWR, and the difference of EE between FD-TWR and HD-TWR grows quickly, as the imbalance increases. We also see that FD-TWR-2TS is more robust against the traffic imbalance, compared to FD-TWR-1TS. Particularly, Fig. 6(a) shows consistently across the spectrum of PA efficiency that FD-TWR-1TS is suitable for balanced traffic, while FD-TWR-2TS can provide higher EE in the case of severely unbalanced traffic. Fig. 6(b) confirms this with consistent results across a wide spectrum of self-cancellation capability. As given earlier, the reason is that FD-TWR-1TS needs to explicitly align the transmission durations of all the nodes. Unbalanced traffic results in excessively high transmit power requirement in some of the links and energy-inefficiently, excessively long transmit duration in the other links, thus compromising the EE.

In addition, Fig. 6(c) reveals that the PA efficiency has a dominant effect over the self-cancellation capability, in the case of balanced traffic. This is because a high level of self-cancellation can reduce the the transmit power requirement which is only part of the total PA consumption. In contrast, increasing the PA efficiency can reduce the total PA consumption effectively, thereby resulting in more noticeable EE improvement.

VII. CONCLUSION

In this paper, the EE maximization of FD-TWR is formulated under practical non-ideal PAs and non-negligible transmission-dependent circuit power. Reformulations of the problems are carried out by deriving the necessary conditions of the optimal solutions, and then rigorously proved to be convex. Extensive simulations show the superiority of FD-TWR in terms of EE when traffic demand is high. The simulations also reveal that the maximum EE of FD-TWR is more sensitive to the PA efficiency than it is to self-cancellation. Also, a full FD design of FD-TWR, i.e., FD-TWR-1TS, is susceptible to unbalanced traffic, while the design with only the relay operating in FD mode, i.e., FD-TWR-2TS, exhibits strong tolerance to traffic imbalance.
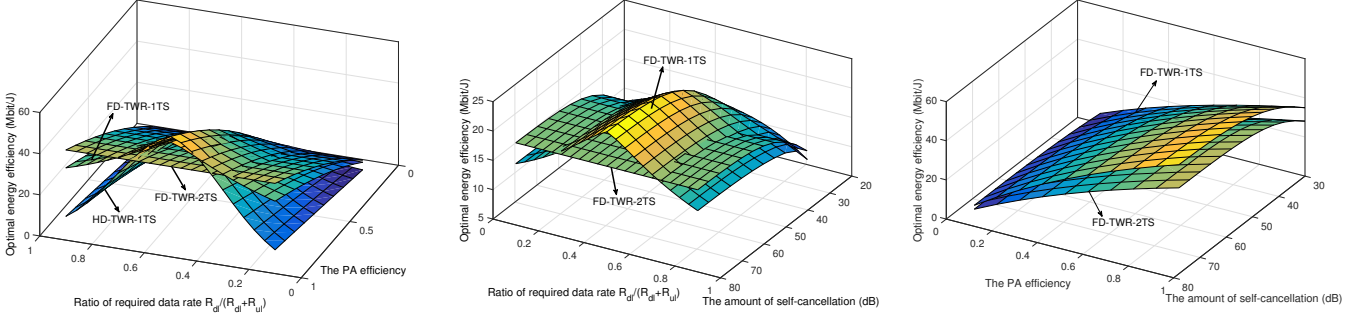
APPENDIX A
PROOF OF THE NECESSARY CONDITION OF (P1)

Proof: We can rewrite the rate constraints in (P1), as given by

$$C_{a,r} = \frac{t_1}{T} W \log_2(1 + \frac{P_a|h_{a,r}|^2}{P_{r,b}\left|\tilde{h}_{r,r}\right|^2 + \sigma_r^2}) \geq R_{fl}, \quad (44)$$

$$C_{r,b} = \frac{t_1}{T} W \log_2(1 + \frac{P_{r,b}|h_{r,b}|^2}{\sigma_b^2}) \geq R_{fl}, \quad (45)$$

(a) The impact of PA efficiency and traffic imbalance where PS ($\alpha_{PS} = 40$ dB) is adopted.

(b) The impact of traffic imbalance and self-cancellation.

(c) The impact of PA efficiency and self-cancellation under symmetric traffic.

Fig. 6. The impact of PA efficiency, self-cancellation and traffic imbalance under ETPA model where $R_{fl} + R_{rl} = 60$ Mbps.

$$C_{b,r} = \frac{t_2}{T}W\log_2(1 + \frac{P_b|h_{b,r}|^2}{P_{r,a}|\tilde{h}_{r,r}|^2 + \sigma_r^2}) \geq R_{rl}, \quad (46)$$

$$C_{r,a} = \frac{t_2}{T}W\log_2(1 + \frac{P_{r,a}|h_{r,a}|^2}{\sigma_a^2}) \geq R_{rl}. \quad (47)$$

Assume that (**P1**) is feasible, i.e., $P_{\max,i}$, $i \in \{a, b, r\}$, and/or T are sufficiently large to accommodate R_{fl} and R_{rl}. We confirm that these constraints are all active at the optimal solution for (**P1**), i.e., equalities are taken in all the constraints. This can be proved by assuming the optimal solution, denoted by $\{P_a^*, P_b^*, P_{r,a}^*, P_{r,b}^*, t_1^*, t_2^*\}$, is inside the feasible solution region where $0 \leq P_i^* \leq \Psi^{-1}(P_{\max,i})$ and $0 \leq P_{r,i}^* \leq \Psi^{-1}(P_{\max,r})$ for $i \in \{a, b\}$; $0 \leq t_1^* + t_2^* \leq T$; and some or all of the constraints (44)–(47) are inactive.

Under this assumption, given t_1^* and t_2^* ($t_1^* + t_2^* \leq T$), we can reduce $P_{r,b}^*$ to $P_{r,b}^*{}'$ and $P_{r,a}^*$ to $P_{r,a}^*{}'$ until equalities can be taken in (45) and (47), respectively. We can proceed to reduce P_a^* to $P_a^*{}'$ and P_b^* to $P_b^*{}'$ until equalities are taken in (44) and (46), respectively. As a result, the total energy consumption is reduced. Clearly, $0 \leq P_a^*{}' < P_a^*$, $0 \leq P_b^*{}' < P_b^*$, $0 \leq P_{r,a}^*{}' < P_{r,a}^*$, and $0 \leq P_{r,b}^*{}' < P_{r,b}^*$; in other words, $\{P_a^*{}', P_b^*{}', P_{r,a}^*{}', P_{r,b}^*{}', t_1^*, t_2^*\}$ is also a feasible solution for (**P1**). This contradicts with the assumption that $\{P_a^*, P_b^*, P_{r,a}^*, P_{r,b}^*, t_1^*, t_2^*\}$ is the optimal solution. The necessary conditions of the optimal solution for (**P1**) are proved.

As a matter of fact, given any pair of feasible t_1 and t_2 ($t_1 + t_2 \leq T$), P_a, P_b, $P_{r,a}$ and $P_{r,b}$ are all minimized (and so is the total energy consumption) if and only if (44)–(47) all take equalities. The minimized powers all satisfy the rate constraints (44)–(47), but may violate the power constraints of $P_{\max,i}$. In this case, we adjust t_1 and t_2, until the powers satisfy the power constraints and the total energy consumption is minimized over $\{t_1, t_2\}$, as described in Section IV. It is possible that (**P1**) is infeasible, that is, for any possible t_1 and t_2, the minimized powers violate the power constraints; in other words, by no means can the system of interest meet the data rate requirements R_{fl} and/or R_{rl}. The data rate requirements need to be reduced, which is beyond the scope of this paper. ∎

APPENDIX B
PROOF OF THE NECESSARY CONDITION OF (**P2**)

Proof: The rate constraints in (**P2**) can be rewritten as

$$C_{a,r} = \frac{t_1}{T}W\log_2\left(\frac{P_a|h_{a,r}|^2}{P_a|h_{a,r}|^2 + P_b|h_{b,r}|^2} + \frac{P_a|h_{a,r}|^2}{P_r|\tilde{h}_{r,r}|^2 + \sigma_r^2}\right) \geq R_{fl}, \quad (48)$$

$$C_{b,r} = \frac{t_1}{T}W\log_2\left(\frac{P_b|h_{b,r}|^2}{P_a|h_{a,r}|^2 + P_b|h_{b,r}|^2} + \frac{P_b|h_{b,r}|^2}{P_r|\tilde{h}_{r,r}|^2 + \sigma_r^2}\right) \geq R_{rl}, \quad (49)$$

$$C_{r,a} = \frac{t_1}{T}W\log_2(1 + \frac{P_r|h_{r,a}|^2}{P_a|\tilde{h}_{a,a}|^2 + \sigma_a^2}) \geq R_{rl}, \text{ or} \quad (50a)$$

$$C_{r,b} = \frac{t_1}{T}W\log_2(1 + \frac{P_r|h_{r,b}|^2}{P_b|\tilde{h}_{b,b}|^2 + \sigma_b^2}) \geq R_{fl}. \quad (50b)$$

Assume that (**P2**) is feasible, i.e., $P_{\max,i}$, $i \in \{a, b, r\}$, and/or T are sufficiently large to accommodate R_{fl} and R_{rl}. We confirm that both (48) and (49) are active, and at least one of (50a) and (50b) is active, at the optimal solution for (**P2**). This can be proved by assuming that the optimal solution, denoted by $\{P_a^*, P_b^*, P_r^*, t_1^*\}$, is inside the feasible solution region where $0 \leq P_i^* \leq \Psi^{-1}(P_{\max,i})$ for $i \in \{a, b, r\}$; $0 \leq t_1^* \leq T$; and some or all of the constraints (48)–(50) are inactive.

Under this assumption, given t_1^* and P_r^*, we certainly can reduce P_a^* and P_b^* in an alternating manner until equalities are taken in both (48) and (49). Specifically, fixing P_b^* and reducing P_a^* leads to a reduction of $C_{a,r}$ in the left-hand side (LHS) of (48) and a growth of $C_{b,r}$ in the LHS of (49); while fixing P_a^* and reducing P_b^* leads to a reduction of $C_{b,r}$ and a growth of $C_{a,r}$. Nevertheless, this process is surely convergent upon both (48) and (49) taking the equalities, since P_a^* and P_b^* are not only strictly monotonically decreasing but also lower bounded by

$$P_a^* \geq \left(2^{\frac{R_{rl}T}{Wt_1^*}} - \frac{P_a^*|h_{a,r}|^2}{P_a^*|h_{a,r}|^2 + P_b^*|h_{b,r}|^2}\right)\frac{P_r^*|\tilde{h}_{r,r}|^2 + \sigma_r^2}{|h_{a,r}|^2} \quad (51)$$

$$> \left(2^{\frac{R_{rl}T}{Wt_1^*}} - 1\right)\frac{\sigma_r^2}{|h_{a,r}|^2} > 0,$$

$$P_b^* \geq \left(2^{\frac{R_{fl}T}{Wt_1^*}} - \frac{P_b^*|h_{b,r}|^2}{P_a^*|h_{a,r}|^2 + P_b^*|h_{b,r}|^2}\right) \frac{P_r^*\left|\tilde{h}_{r,r}\right|^2 + \sigma_r^2}{|h_{b,r}|^2} \quad (52)$$

$$> \left(2^{\frac{R_{fl}T}{Wt_1^*}} - 1\right) \frac{\sigma_r^2}{|h_{b,r}|^2} > 0.$$

Given P_a^* and P_b^* taking equalities in (48) and (49), we can proceed to reduce P_r^* until either of (50a) and (50b) takes equality. Both $C_{a,r}$ and $C_{b,r}$ on the LHS of (48) and (49) grow. Given P_r^*, we continue to reduce P_a^* and P_b^* till (48) and (49) take equalities again, as described above. This process is surely convergent, i.e., (48), (49), and (50a) or (50b), can all take equalities, since P_r^* is not only strictly monotonically decreasing but also lower bounded by

$$P_r^* \geq \max\left\{(2^{\frac{R_{rl}T}{Wt_1^*}} - 1)\frac{P_a^*\left|\tilde{h}_{a,a}\right|^2 + \sigma_a^2}{|h_{r,a}|^2}; (2^{\frac{R_{fl}T}{Wt_1^*}} - 1)\frac{P_b^*\left|\tilde{h}_{b,b}\right|^2 + \sigma_b^2}{|h_{r,b}|^2}\right\}$$

$$> \max\left\{(2^{\frac{R_{rl}T}{Wt_1^*}} - 1)\frac{\sigma_a^2}{|h_{r,a}|^2}; (2^{\frac{R_{fl}T}{Wt_1^*}} - 1)\frac{\sigma_b^2}{|h_{r,b}|^2}\right\} > 0. \quad (53)$$

As a result, all P_a^*, P_b^* and P_r^* can be continually reduced to a fixed point where both (48) and (49) are active and at least one of (50a) and (50b) is active. Clearly, the fixed point is a feasible solution for (**P2**), and requires less energy than the assumed optimal solution. This contradicts with the assumption that the assumed solution $\{P_a^*, P_b^*, P_r^*, t_1^*\}$ is the optimal feasible solution for (**P2**). The necessary conditions of the optimal solution for (**P2**) are proved.

We note that, given any $t_1 \leq T$, P_a, P_b and P_r are all minimized (and so is the total energy consumption) if and only if (48)–(50) are all active. The minimized powers all satisfy the rate constraints (48)–(50), but may violate the power constraints of $P_{max,i}$. In this case, we adjust t_1, until the powers satisfy the power constraints and the total energy consumption is minimized over $t_1 \in [0,T]$, as described in Section V. It is possible that (**P2**) is infeasible, that is, for any possible t_1, the minimized powers violate the power constraints; in other words, by no means can the system of interest meet the data rate requirements R_{fl} and/or R_{rl}. The data rate requirements need to be reduced. ∎

APPENDIX C
ENERGY EFFICIENCY OF HD-TWR-2TS

PNC is also adopted in HD-TWR and two timeslots are required to accomplish TWR. During the first timeslot, nodes A and B transmit signal to node R simultaneously. The received signal at node R can be given by

$$y_r = \sqrt{P_a}\,h_{a,r}\,x_a + \sqrt{P_b}\,h_{b,r}\,x_b + n_r. \quad (54)$$

With nested lattice code, the achievable data rates satisfy

$$C_{a,r} = \frac{t_1}{T}W\log_2\left(\frac{P_a|h_{a,r}|^2}{P_a|h_{a,r}|^2 + P_b|h_{b,r}|^2} + \frac{P_a|h_{a,r}|^2}{\sigma_r^2}\right), \quad (55)$$

$$C_{b,r} = \frac{t_1}{T}W\log_2\left(\frac{P_b|h_{b,r}|^2}{P_a|h_{a,r}|^2 + P_b|h_{b,r}|^2} + \frac{P_b|h_{b,r}|^2}{\sigma_r^2}\right). \quad (56)$$

During the second timeslot, node R forwards the combined signal x_r to nodes A and B. Thus, the received signal at nodes A and B can be given by

$$y_i = \sqrt{P_r}h_{r,i}x_r + n_i, \ i \in \{a,b\}. \quad (57)$$

By cancelling the self-interfere, the achievable data rates at nodes A and B in the second timeslot satisfy [21]

$$C_{r,a} = \frac{t_2}{T}W\log_2(1 + \frac{P_r|h_{r,a}|^2}{\sigma_a^2}), \quad (58)$$

$$C_{r,b} = \frac{t_2}{T}W\log_2(1 + \frac{P_r|h_{r,b}|^2}{\sigma_b^2}). \quad (59)$$

The EE maximization problem has the same form as that of FD-TWR-2TS, but the total energy consumption is different, as given by

$$E_{\text{total}} = (P_{tx,a} + P_{tx,b} + P_{rx,r})\,t_1 \quad (60)$$
$$+ (P_{rx,a} + P_{rx,b} + P_{tx,r})\,t_2 + P_{idle}\,(T - t_1 - t_2),$$

where $P_{idle} = P_{idle,a} + P_{idle,b} + P_{idle,r}$. The necessary conditions can also be given by (30) and (31), as can be proved in the same way as FD-TWR-1TS. The optimal transmit powers, P_a, P_b and P_r, are written as

$$P_a(t_1) = \left(\lambda_1 - \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)\frac{\sigma_r^2}{|h_{a,r}|^2}, \quad (61)$$

$$P_b(t_1) = \left(\lambda_2 - \frac{\lambda_2}{\lambda_1 + \lambda_2}\right)\frac{\sigma_r^2}{|h_{b,r}|^2}, \quad (62)$$

$$P_r(t_2) = \max\left\{(\lambda_3 - 1)\frac{\sigma_b^2}{|h_{r,b}|^2}, (\lambda_4 - 1)\frac{\sigma_a^2}{|h_{r,a}|^2}\right\}, \quad (63)$$

where $\lambda_1 = 2^{\frac{R_{fl}T}{Wt_1}}$, $\lambda_2 = 2^{\frac{R_{rl}T}{Wt_1}}$, $\lambda_3 = 2^{\frac{R_{fl}T}{Wt_2}}$ and $\lambda_4 = 2^{\frac{R_{rl}T}{Wt_2}}$ for notational simplicity.

By substituting (61)–(63) into the optimization of HD-TWR-2TS, we can obtain the equivalent objective function under the TPA model, as given by

$$E(t_1, t_2) = [\sqrt{\frac{P_{max,a}}{\eta_{max,a}^2}}\sqrt{P_a(t_1)} + \sqrt{\frac{P_{max,b}}{\eta_{max,b}^2}}\sqrt{P_b(t_1)} + P_1]t_1$$

$$+ [\sqrt{\frac{P_{max,r}}{\eta_{max,r}^2}}\sqrt{P_r(t_2)} + P_2]t_2$$

$$+ (P_{idle,a} + P_{idle,b} + P_{idle,r})T,$$

where $P_1 = P_{rx,r} + P_{base,a} + P_{base,b} + \varepsilon(R_{fl} + R_{rl}) - P_{idle}$, $P_2 = P_{rx,a} + P_{rx,b} + P_{base,r} + \varepsilon\max\{R_{fl}, R_{rl}\} - P_{idle}$; and the equivalent objective function under the ETPA model, as formulated by

$$E(t_1, t_2) = [\frac{P_a(t_1)}{(1 + u\kappa_a)\eta_{max,a}} + \frac{P_b(t_1)}{(1 + u\kappa_b)\eta_{max,b}} + P_1]t_1$$

$$+ [\frac{P_r(t_2)}{(1 + u\kappa_r)\eta_{max,r}} + P_2]t_2 + P_{idle}T,$$

where $P_1 = \frac{u\kappa_a P_{max,a}}{(1 + u\kappa_a)\eta_{max,a}} + \frac{u\kappa_b P_{max,b}}{(1 + u\kappa_b)\eta_{max,b}} + P_{rx,r} + P_{base,a} + P_{base,b} + \varepsilon(R_{fl} + R_{rl}) - P_{idle}$ and $P_2 = \frac{u\kappa_r P_{max,r}}{(1 + u\kappa_r)\eta_{max,r}} + P_{rx,a} + P_{rx,b} + P_{base,r} + \varepsilon\max\{R_{fl}, R_{rl}\} - P_{idle}$.

It can be proved that $E(t_1, t_2)$ is quasi-convex under the TPA model and convex under the ETPA model. The detailed proofs are suppressed due to the limited space.

REFERENCES

[1] G. Liu, F. Yu, and H. Ji *et al.*, "In-band full-duplex relaying: A survey, research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 500-524, Jan. 2015.

[2] Z. Wei, X. Zhu, and S. Sun *et al.*, "Full-duplex vs. half-duplex amplify-and-forward relaying: which is more energy efficient in 60 GHz dual-hop indoor wireless systems?," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2936-2947, Dec. 2015.

[3] A. Sabharwal, P. Schniter, and D. Guo *et al.*, "In-band full-duplex wireless: challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637-1652, Sep. 2014.

[4] E. Ahmed, A. M. Eltawil and A. Sabharwal, "Rate gain region and design tradeoffs for full-duplex wireless communications," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3556-3565, Jul. 2013.

[5] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296-4307, Dec. 2012.

[6] Y. Liu, L. Lu, G. Y. Li *et al.*, "Joint user association and spectrum allocation for small cell networks with wireless backhauls," *IEEE Wireless Commun. Lett.*, vol. 5, no. 5, pp. 496-499, Oct. 2016.

[7] B. Rankov, and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE ISIT*, pp. 1668-1672, Jul. 2006.

[8] M. P. Wilson, K. Narayanan, and H. D. Pfister *et al.*, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5641-5654, Nov. 2010.

[9] H. Kim, N. Lee, and J. Kang, "Energy efficient two-way AF relay system with multiple-antennas," in *Proc. IEEE VTC-Fall*, pp. 1-5, Sep. 2011.

[10] M. Zhou, Q. Cui, and R. Jäntti *et al.*, "Energy-efficient relay selection and power allocation for two-way relay channel with analog network coding," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 816-819, Jun. 2012.

[11] F. Parzysz, M. Vu, and F. Gagnon, "Energy minimization for the half-duplex relay channel with decode-forward relaying," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2232-2247, Jun. 2013.

[12] C. Sun, and C. Yang, "Energy-efficient hybrid one-way and two-way relay transmission," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 3737-3751, Oct. 2013.

[13] M. Hossain, K. Koufos, and R. Jäntti, "Minimum-energy power and rate control for fair scheduling in the cellular downlink under flow level delay constraint," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3253-3263, Jul. 2013.

[14] S. Cripps, "RF power amplifiers for wireless communications," 2nd ed. Norwood, MA: Artech House, 2006.

[15] Q. Cui, T. Yuan, and W. Ni, "Energy-efficient two-way relaying under non-ideal power amplifiers," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1257-1270, Feb. 2017.

[16] S. Kim, and Y. Lee, "Energy-efficient power allocation for OFDM signaling over a two-way AF relay," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4856-4863, Nov. 2014.

[17] C. Xiong, L. Lu, and G. Y. Li, "Energy-efficient OFDMA-based two-way relay," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3157-3169, Sep. 2015.

[18] Q. Cui, T. Yuan, and X. Tao *et al.*, "Energy efficiency analysis of two-way DF relay system with non-ideal power amplifiers," *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1254-1257, Jul. 2014.

[19] H. Chen, G. Li, and J. Cai, "Spectral-energy efficiency tradeoff in full-duplex two-way relay networks," to appear *IEEE Syst. J.*, 2017.

[20] Z. Zhang, Z. Chen, and M. Shen *et al.*, "Spectral and energy efficiency of multi-pair two-way full-duplex relay systems with massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 848-863, Apr. 2016.

[21] W. Nam, S. Chung, and Y. Lee, "Capacity of the Gaussian two-way relay channel to within 1/2 bit," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5488-5494, Nov. 2010.

[22] X. Jia, P. Deng, and L. Yang *et al.*, "Spectrum and energy efficiencies for multiuser pairs massive MIMO systems with full-duplex amplify-and-forward relay," *IEEE Access*, vol. 3, pp. 1907-1918, Oct. 2015.

[23] G. Liu, F. Yu, and H. Ji *et al.*, "Energy-efficient resource allocation in cellular networks with shared full-duplex relaying," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3711-3724, Aug. 2015.

[24] M. Khafagy, A. Ismail, and M. Alouini *et al.*, "Energy-efficient co-operative protocols for full-duplex relay channels," in *Proc. IEEE GLOBECOM*, pp. 362-367, Dec. 2013.

[25] N. Zlatanov, and R. Schober, "Buffer-aided half-duplex relaying can outperform ideal full-duplex relaying," *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 479-482, Mar. 2013.

[26] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 545-554, Feb. 2010.

[27] X. Wang, and Z. Li, "Energy-efficient transmissions of bursty data packets with strict deadlines over time-varying wireless channels," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2533-2543, May 2013.

[28] Z. Nan, X. Wang, and W. Ni, "Energy-efficient transmission of delay limited bursty data packets under non-ideal circuit power consumption." in *Proc. IEEE ICC*, pp. 4957-4962, Jun. 2014.

[29] X. Chen, W. Ni, and X. Wang *et al.*, "Provisioning quality-of-service to energy harvesting wireless communications." *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 102-109, Apr. 2015.

[30] W. Li, J. Lilleberg, and K. Rikkinen, "On rate region analysis of half-and full-duplex OFDM communication links," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1688-1698, Sep. 2014.

[31] W. Ni, I. B. Collings, and X. Wang *et al.*, "Multi-hop point-to-point FDD wireless backhaul for mobile small cells," *IEEE Wireless Commun.*, vol. 21, no. 4, pp. 88-96, Aug. 2014.

[32] W. Ni, and I. B. Collings, "Adaptive adjacent-frequency interference mitigation in multi-hop point-to-point FDD wireless backhaul networks," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 1988-1991, Dec. 2012.

[33] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074-3085, Sep. 2011.

[34] H. Cui, M. Ma, and L. Song *et al.*, "Relay selection for two-way full duplex relay networks with amplify-and-forward protocol," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3768-3777, Jul. 2014.

[35] K. Nasr, J. Cosmas, and M. Bard *et al.*, "Performance of an echo canceller and channel estimator for on-channel repeaters in DVBT/H networks," *IEEE Trans. Broadcast.*, vol. 53, no. 3, pp. 609-618, Sep. 2007.

[36] C. Xiong, G. Li, and S. Zhang *et al.*, "Energy- and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3874-3886, Nov. 2011.

[37] R. D. Buda, "The upper error bound of a new near-optimal code," *IEEE Trans. Inf. Theory*, vol. 21, no. 4, pp. 441-445, Jul. 1975.

[38] R. D. Buda, "Some optimal codes have structure," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 6, pp. 893-899, Aug. 1989.

[39] U. Erez, and R. Zamir, "Achieving 1/2 log (1+SNR) on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293-2314, Oct. 2004.

[40] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349-2360, Sep. 2005.

[41] G. Auer, O Blume, and V Giannini, "D2.3: Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," *EARTH INFSO-ICT*, ver. 2.0, 2012.

[42] 3GPP TR 36.104, "Base Station (BS) radio transmission and reception," v11.8.2, Apr. 2014. [online] http://www.3gpp.org.

Qimei Cui (M'09-SM'15) (cuiqimei@bupt.edu.cn) received the B.E. and M.S. degrees in Electronic Engineering from Hunan University, Changsha, China, in 2000 and 2003, respectively. She received her Ph.D. degree in Information and Communications Engineering from Beijing University of Posts and Telecommunications (BUPT), Beijing, China in 2006. She was a full professor in School of Information and Communication Engineering at BUPT in 2016. She was guest professor in Department of Electronic Engineering of University of Notre Dame in the United States in 2016. Her main research interests include spectral-efficiency or energy-efficiency based transmission theory and networking technology for 4G/5G broadband wireless communications and green communications. She serves as guest editor of EURASIP Journal on Wireless Communications and Networking, Technical Program Committee (TPC) Member of some International Conferences such as IEEE ICC, IEEE WCNC, IEEE PIMRC, IEEE ICCC, IEEE WCSP 2013, IEEE ISCIT 2012 etc. She was once awarded the only Best Paper Award at IEEE ISCIT 2012, Best Paper Award in IEEE WCNC 2014, the Honorable Mention Demo Award at ACM MobiCom 2009, and the Young Scientist Award at URSI GASS 2014 etc.

Yuhao Zhang received the B.E. degree in communication engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2015, where he is currently pursuing the M.S. degree in information and communication engineering. His research interests include energy efficiency of wireless networks and cooperative communications.

Riku Jäntti (M'02-SM'07) received the M.Sc degree (with distinction) in electrical engineering in 1997 and the D.Sc degree (with distinction) in automation and systems technology in 2001, both from Helsinki University of Technology (TKK). Currently, he is an Associate Professor (tenured) in communications engineering and the Head of the Department of Communications and Networking at Aalto University School of Electrical Engineering, Finland. Prior to joining Aalto (formerly known as TKK) in August 2006, he was Professor Pro Tem with the Department of Computer Science, University of Vaasa. His research interests include radio resource control and optimization for machine type communications, cloud based radio access networks, spectrum and coexistence management and RF inference. He is a senior member of IEEE and Associate Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He is also an IEEE VTS Distinguished Lecturer (Class 2016).
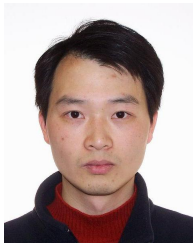
Wei Ni (M'09-SM'15) received the B.E. and Ph.D. degrees in Electronic Engineering from Fudan University, Shanghai, China, in 2000 and 2005, respectively. Currently he is a Senior Scientist, Team and Project Leaders at CSIRO, Australia. He also holds adjunct positions at the University of New South Wales (UNSW), Macquarie University (MQ) and the University of Technology Sydney (UTS). Prior to this he was a postdoctoral research fellow at Shanghai Jiaotong University (2005-2008), Deputy Project Manager at the Bell Labs R&I Center, Alcatel/Alcatel-Lucent (2005-2008), and Senior Researcher at Devices R&D, Nokia (2008-2009). His research interests include stochastic optimization, game theory, graph theory, as well as their applications to network and security. Dr Ni serves as Editor for Hindawi Journal of Engineering since 2012, secretary of IEEE NSW VTS Chapter since 2015, Track Chair for VTC-Spring 2017, Track Co-chair for IEEE VTC-Spring 2016, and Publication Chair for BodyNet 2015. He also served as Student Travel Grant Chair for WPMC 2014, a Program Committee Member of CHINACOM 2014, a TPC member of IEEE ICC14, ICCC15, EICE14, and WCNC10.

Mikko Valkama (S'00-M'02-SM'15) was born in Pirkkala, Finland, in 1975. He received the M.Sc. and Ph.D. degrees (Hons.) in electrical engineering from the Tampere University of Technology (TUT), Finland, in 2000 and 2001, respectively. His Ph.D. dissertation was entitled Advanced I/Q Signal Processing for Wideband Receivers: Models and Algorithms. In 2003, he was a visiting PostDoctoral Researcher with the Communications Systems and Signal Processing Institute, San Diego State University, San Diego, CA, USA. He is currently a Full Professor and Laboratory Head with the Laboratory of Electronics and Communications Engineering, TUT. His general research interests include radio communications, communications signal processing, estimation and detection techniques, signal processing algorithms for flexible radios, cognitive radio, full-duplex radio, radio localization, 5G mobile cellular radio networks, digital transmission techniques, such as the different variants of multicarrier modulation methods and OFDM, and radio resource management for ad-hoc and mobile networks. In 2002, he received the Best Ph.D. Thesis Award by the Finnish Academy of Science and Letters.