Prof. Liang Xiao
Editor
IEEE Transactions on Communications

Re: Decision on TCOM-TPS-20-1501 (Intelligent Reflecting Surface-Aided SWIPT: Joint Waveform, Active and Passive Beamforming Design)

Dear Prof. Liang Xiao and Reviewers,

We would like to express our appreciation for the time and effort dedicated to the reviewing of our paper. The manuscript has been revised carefully based on your valuable comments and suggestions. In particular, we clarified the motivations and contributions of this work, added two low-complexity designs with near-optimal rate-energy performance, and considered the cases of imperfect cascaded CSIT and finite IRS reflection states. To answer the questions and concerns, we include a point-to-point response below, which also describes all modifications made to the manuscript. We hope that the revisions are to the satisfaction of the editor and reviewers.

Best Regards,

Yang Zhao, Bruno Clerckx and
Zhenyuan Feng

Reviewer 1

Comment 1.1 — *This paper assumed that the perfect channel state information (CSI) of the whole system is available at the base station. However, since the IRS is in general not equipped with a radio frequency chain, the accurate CSI of the reflecting links established by the IRS is very challenging to obtain. The reviewer wonders if the proposed algorithm can also be applied to a case where the imperfect CSI of the network is available at the base station.*

Response: Thank you for pointing out the CSI acquisition issue regarding the cascaded AP-IRS-UE link. The proposed passive beamforming algorithms rely on such cascaded CSIT, and recent researches proposed element-wise on/off switching [25], joint training sequence and reflection pattern design [R1] and compressed sensing [R2] techniques to solve this issue. Those references have been cited in the literature review. To address the comment, we consider an imperfect CSIT model where the estimation of the cascaded link at subband n is

$$\hat{\boldsymbol{V}}_n = \boldsymbol{V}_n + \tilde{\boldsymbol{V}}_n \tag{1}$$

where $\tilde{\boldsymbol{V}}_n$ is the estimation error with entries following i.i.d. CSCG distribution $\mathcal{CN}(0, \epsilon_n^2)$. Simulation result in Fig. 1 demonstrates the robustness of the proposed passive beamforming algorithm to cascaded CSIT inaccuracy for broadband SWIPT with different number of IRS elements.

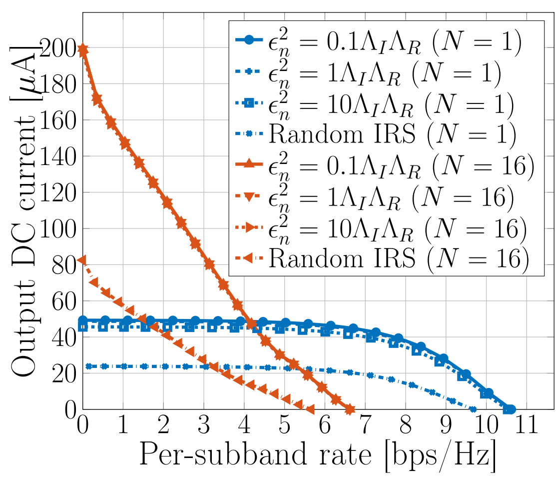


Figure 1: Average R-E region with imperfect cascaded CSIT for $M = 1$, $N = 16$, $L = 20$, $\sigma_n^2 = -40\,\text{dBm}$, $B = 10\,\text{MHz}$ and $d_\text{H} = d_\text{V} = 2\,\text{m}$.

Comment 1.2 — *This paper considered a relatively simple system model where there is only one user in the system. However, in practice, there can be many users co-existing in the SWIPT systems and there may also be two different and independent quality of service requirements: energy harvesting requirement and information decoding requirement. It could be better if the authors can clarify if the proposed algorithm can be employed in a more general case where the base station and the IRS cooperate to serve multiple users in the same time slot.*

Response: Multi-user IRS-aided SWIPT is very interesting and should be investigated in the future, but is left for future work for the reasons explained below.

First, it is important to remind the reviewer that the problem of joint waveform and beamforming design for a multi-user SWIPT system without IRS has never been studied in the literature. In other words, it would not make sense at this stage to address this multiuser IRS-aided SWIPT since the underlying building block of multiuser SWIPT is not known. Only once the problem of multiuser joint waveform and beamforming for multi-user SWIPT has been addressed, can we investigate the multiuser joint waveform and beamforming design for IRS-aided SWIPT. Those problems could be addressed in future research, but are not the scope of this paper.

Second, the emphasis is on single-user because we believe a proper understanding of the single-user case is crucial before jumping into multi-user scenarios. Our modeling of SWIPT is not "simple": not only we transmit both information and power, we deal with joint space and frequency optimization through a joint beamforming and waveform problem that is challenging due to the nonlinearity of the rectifier that induces coupling among frequencies components. This completely contrasts with any existing IRS-aided SWIPT work as discussed in the introduction [31 – 33].

Third, multi-user WPT leads to its own set of challenges compared single-user WPT. Taking an non-IRS-aided WPT, frequency and spatial domains can be decoupled in single-user case, while they cannot be decoupled in the multi-user case, as shown in [8, 36]. Consequently, single-user and multi-user WPT are designed differently in [8, 36]. Though the SWIPT case has not been studied, the same set of problems is expected to occur between single user and multiuser case and would be even more challenging since communication is delivered simultaneously with power. Adding IRS into the picture would make it even more complicated.

Comment 1.3 — *It is well known that after employing semidefinite programming for handling the phase shift matrix at the IRS, it is very unlikely to obtain a rank-one phase shift matrix without any further modification. The reviewer notices that the authors proposed the Gaussian randomization method to ensure a rank-one solution. Also, the convergence of the proposed overall algorithm also relies on the unit-rank solution. Therefore, to make this paper more comprehensive and convincing, it is suggested to provide more results (such as figures, tables, and data analysis) in the simulation part to show that the rank-one solution can always be obtained even without applying the Gaussian randomization. As this is a very important and interesting conclusion to the colleagues working in the same area, it could be better if the authors can further discuss, interpret, and clarify this in a remark.*

Response: Thank you for raising this issue. The rank-1 property of the IRS matrix $\mathbf{\Phi}$ to the relaxed problem (25a) – (25d) is suggested by the CVX software [39] under different configurations. Although the strict convergence and local optimality of the BCD Algorithm 3 relies on this conclusion, it is intricate to provide a mathematical proof due to the coupled objective (24) and the log sum constraint (25b). Instead, we define the eigen ratio as $\nu = \max_l \lambda_l(\mathbf{\Phi}) / \sum_l \lambda_l(\mathbf{\Phi})$ where $\lambda_l(\mathbf{\Phi})$ is the l-th eigenvalue of $\mathbf{\Phi}$. Since $\nu = 1$ means $\mathbf{\Phi}$ is rank-1, we claim $\mathbf{\Phi}$ is approximately rank-1 when $1 - \nu \approx 0$. Table 1 shows the maximum value of $1 - \nu$ over 20 R-E samples of 300 tested channel realizations for both BCD algorithms. With reasonable precision, we conclude that $\mathbf{\Phi}$ is always rank-1 under different configurations for all tested channel realizations. It suggests that the rank constraint (25e) can be relaxed with negligible performance loss, $\bar{\boldsymbol{\phi}}$ can be directly obtained through eigen decomposition, and the assumption of Proposition 3 always hold such that the BCD Algorithm 3 converges to local optimal points.

Table 1: The maximum value of $1 - \nu$ over all R-E samples during BCD and LC-BCD algorithms for all tested channel realizations

		M			N			L			B	
		1	2	4	1	4	16	20	40	80	1 MHz	10 MHz
$\max(1-\nu)$ $[\times 10^{-16}]$	BCD	5.5511	6.6613	6.6613	5.5511	6.6613	5.5511	6.6613	5.5511	6.6613	6.6613	5.5511
	LC	6.6613	6.6613	6.6613	5.5511	6.6613	5.5511	6.6613	8.8818	6.6613	6.6613	5.5511

Comment 1.4 — *The figures in the current version are relatively small, it is suggested to provide larger figures to help the readers better understand the results.*

Response: We sincerely apologize for the inconvenience. The figures would be amplified in future versions.

Comment 1.5 — *In the current version, the authors claimed that by applying semidefinite relaxation and omitting the rank-one constraint, the performance loss is negligible. The reviewer wonders if this is because of the relatively simple system model, as there is only one single user who has both power and information requirements. It could be better if the authors can further discuss this issue for a more general multiple user scenario.*

Response: We agree with the reviewer that the relaxation on the rank-one constraint deserves more attention. However, as discussed in Response 1.2, multi-user IRS-aided SWIPT is out the scope of this paper and we would like to provide some general ideas to address this issue. For any passive beamforming problem with unit-rank constraint (which is non-convex), we can either relax the constraint and use Gaussian randomization method to obtain a high-quality solution [38], or replace the unit-rank constraint $\text{rank}(\boldsymbol{\Phi}) = 1$ with the constraint on largest singular value $\text{Tr}(\boldsymbol{\Phi}) - \sigma_{\max}(\boldsymbol{\Phi}) \leq 0$ then solve the problem by the penalty method [33], [R3]. Both techniques are expected to provide a unit-rank solution with performance very close to the original solution.

Reviewer 2

Comment 2.1 — *New RIS models are now being adopted as in [24] where it has been shown that the reflected signals depend on the direction of the arriving signal and this needs to be included in the analysis for realistic quantification.*

Response: Thank you for sharing this paper. It investigated the impact of non-zero effective resistance on the reflection pattern and pointed out that the amplitude of the reflection coefficient depends on the phase shift forced on the incoming signal when power dissipation is considered at the IRS. It also proposed an analytical IRS model together with an BCD algorithm to maximize the achievable rate by passive beamforming. Simulation results emphasized the importance of modeling such a relationship in practical IRS design. There exist various refined models of IRS in the literature, however, we have finally decided to use the most common and simplest IRS model at the current stage to reduce the design complexity and provide a primary benchmark for practical IRS-aided SWIPT.

Comment 2.2 — *Why is MRT considered as precoder by (27) rather than optimizing it? Is it globally optimal too?*

Response: In the single-user scenario, the global optimal information and power precoders coincide at MRT. To prove this, we decouple the waveform in the spatial and frequency domains by

$$\boldsymbol{w}_{\text{I/P},n} = s_{\text{I/P},n}\boldsymbol{b}_{\text{I/P},n} \tag{2}$$

where $s_{\text{I/P},n}$ denotes the amplitude of modulated/multisine waveform at tone n, and $\boldsymbol{b}_{\text{I/P},n}$ denotes the information/power precoder. The MRT precoder at subband n is given by

$$\boldsymbol{b}^{\star}_{\text{I/P},n} = \frac{\boldsymbol{h}_n}{\|\boldsymbol{h}_n\|} \tag{3}$$

From the perspective of WIT, the MRT precoder maximizes $|\boldsymbol{h}_n^H \boldsymbol{w}_{\text{I},n}| = \|\boldsymbol{h}_n\| s_{\text{I},n}$ thus maximizes the rate (8). From the perspective of WPT, the MRT precoder maximizes $(\boldsymbol{h}_n^H \boldsymbol{w}_{\text{I/P},n})(\boldsymbol{h}_n^H \boldsymbol{w}_{\text{I/P},n})^* = \|\boldsymbol{h}_n\|^2 s^2_{\text{I/P},n}$ thus maximizes the second and fourth order DC terms (11) – (14). Hence, MRT is the global optimal active precoder and no dedicated energy beams are required in the single-user SWIPT. We have separated the active beamforming design from the waveform design in the revised manuscript to clarify this point.

Comment 2.3 — *Some strong assumptions like perfect CSI availability limit the practical utility of the proposed analytical results.*

Response: The reviewer is referred to Response 1.1. Indeed, the assumption of perfect CSIT is very ideal and the existing channel estimation protocols may not provide decent results in practice. We have investigated the impact of CSIT estimation error of the cascaded AP-IRS-UE link on the R-E performance, and Fig. 1 shows that the proposed algorithms are robust to CSIT inaccuracy.

Comment 2.4 — *All the assumptions and relaxations adopted used in the derivation of results as in (24) need to be explicitly mentioned along with appropriate justification for the same.*

Response: We appreciate your suggestion and have revised the manuscript correspondingly. The original objective function (20) is differentiable and non-concave in $\mathbb{C}^{4N-2}$, and we approximate (linearize) the second-order terms by the first-order Taylor expansions (21) – (23) to formulate SCA problems of the original passive beamforming problem (i.e. maximize (20) s.t. (25b) – (25e)). The objective affine function (24) is obtained by plugging (21) – (23) into (20). It satisfies $\tilde{z}(\mathbf{\Phi}^{(i)}, \mathbf{\Phi}^{(i)}) = z(\mathbf{\Phi}^{(i)}) \geq \tilde{z}(\mathbf{\Phi}^{(i)}, \mathbf{\Phi}^{(i-1)})$ such that solving (25) iteratively is guaranteed to converge to a local optimal point of the original passive beamforming problem.

Comment 2.5 — *Some transformations have been made while solving the original problem, but it has not been explicitly mentioned whether it is equivalent to transformation or not.*

Response: Thank you for the reminder. All transformations are equivalent to their original form and we have mentioned this point explicitly in the revised manuscript.

Comment 2.6 — *Are the proposed solutions locally optimal or globally optimal? It is not clear whether the convergence of proposed solution methodologies is local or global? Also, how fast is it?*

Response: Algorithm 1 – 4 only provide local optimal solutions with analytical/numerical local convergence proof, and the performance indeed depends on the initialization. For the passive beamforming Algorithm 1, we initialize the phase shift of all IRS elements as i.i.d. uniform random variables over $[0, 2\pi)$. For the waveform amplitude and splitting ratio Algorithm 2, we initialize the modulated waveform by the Water-Filling (WF) strategy (37) and the multisine waveform by the Scaled Matched Filter (SMF) scheme (38), and assume $\rho^{(0)} = \bar{\rho}^{(0)} = 1$. These parameters are used for general initialization and regulated by the algorithm afterwards. However, as Algorithm 3 only converge to local optimal solutions, few R-E points might be strictly worse (with less rate and energy) than the others especially when N is very large. To address this issue, we draw the R-E boundary from the high-rate low-energy (lower right) points to the low-rate high-energy (upper left) points. If a point is strictly dominated, we discard the candidate, reinitialize Algorithm 3 by the solution at the previous point, then perform the optimization again. For a tolerance of $\epsilon = 10^{-8}$, Algorithm 1 – 4 typically converge within 2, 7, 2, 2 iterations, respectively.

Comment 2.7 — *The time complexity of the proposed algorithms, especially involving branch and bound methods, seems to be high especially applications assuming perfectly CSI availability as the coherence times are practically pretty low. So, the authors would like to justify it so that the proposed solution can be obtained over relatively short coherence intervals.*

Response: Problem (25) is not a Semidefinite Programming (SDP) due to the $(t_{\mathrm{I},0}^{(i)} - t_{\mathrm{P},0}^{(i)})^2$ term in (24). Hence, existing complexity analysis tools cannot be direct applied to Algorithm 1 and 4. For Algorithm 2, the computational complexity scales exponentially with the number of subbands [36, 43], but an analytical expression cannot be derived on top of relevant literatures. To facilitate practical SWIPT implementation, we propose two closed-form adaptive waveform amplitude design by combining WF and SMF under TS and PS setups. The optimal waveform design for WIT corresponds to the WF strategy that assigns the amplitude of modulated tone n by

$$s_{\mathrm{I},n} = \sqrt{2\left(\mu - \frac{\sigma_n^2}{P\|\boldsymbol{h}_n\|^2}\right)^+} \tag{4}$$

where μ is chosen to satisfy the power constraint $\|\boldsymbol{s}_I\|^2/2 \leq P$. The closed-form solution can be obtained by iterative power allocation [45] and the details are omitted here. On the other hand, SMF was proposed in [12] as a suboptimal WPT resource allocation scheme that assigns the amplitude of sinewave n by

$$s_{\mathrm{P},n} = \sqrt{\frac{2P}{\sum_{n=1}^{N}\|\boldsymbol{h}_n\|^{2\alpha}}}\|\boldsymbol{h}_n\|^{\alpha} \tag{5}$$

where the scaling ratio $\alpha \geq 1$ scales the matched filter to exploit the rectifier nonlinearity. In the low-complexity TS waveform design, modulated waveform (4) is used in the data session while multisine waveform (5) is used in the energy session. In contrast, the low-complexity PS scheme jointly designs the waveform balancing ratio δ and splitting ratio ρ, and the amplitude of modulated and multisine components are given by

$$s_{\mathrm{I},n} = \sqrt{2(1-\delta)\left(\mu - \frac{\sigma_n^2}{P\|\boldsymbol{h}_n\|^2}\right)^+} \tag{6}$$

$$s_{\mathrm{P},n} = \sqrt{\frac{2\delta P}{\sum_{n=1}^{N}\|\boldsymbol{h}_n\|^{2\alpha}}}\|\boldsymbol{h}_n\|^{\alpha} \tag{7}$$

Fig. 2a – 2d demonstrate that the low-complexity BCD algorithm achieves near-optimal performance under different configurations.

Comment 2.8 — *How practical is it to consider lossless reflection from the RIS? Specifically, by considering the magnitude to be 1, the reflection losses at the RIS have been ignored.*

Response: The assumption of lossless reflection is indeed very ideal. Based on the circuit analysis, the authors of [24] shows that the reflection amplitude depends on the phase shift, which experiences a trough at 0 phase shift and is negatively correlated to the element resistance. For $R_l = 1\,\Omega$, $|\phi_l|$ is 1 at a phase shift of π but around 0.6 at a phase shift of 0. From the perspective of field and propagation, [R4] points out that if the dimension of IRS element is much larger than the wavelength, the beamwidth would become small and the scattered field strength would depend on the observation angle. However, to reduce the design complexity and provide a preliminary insight, we considered the commonest lossless reflection pattern at the IRS in this paper.

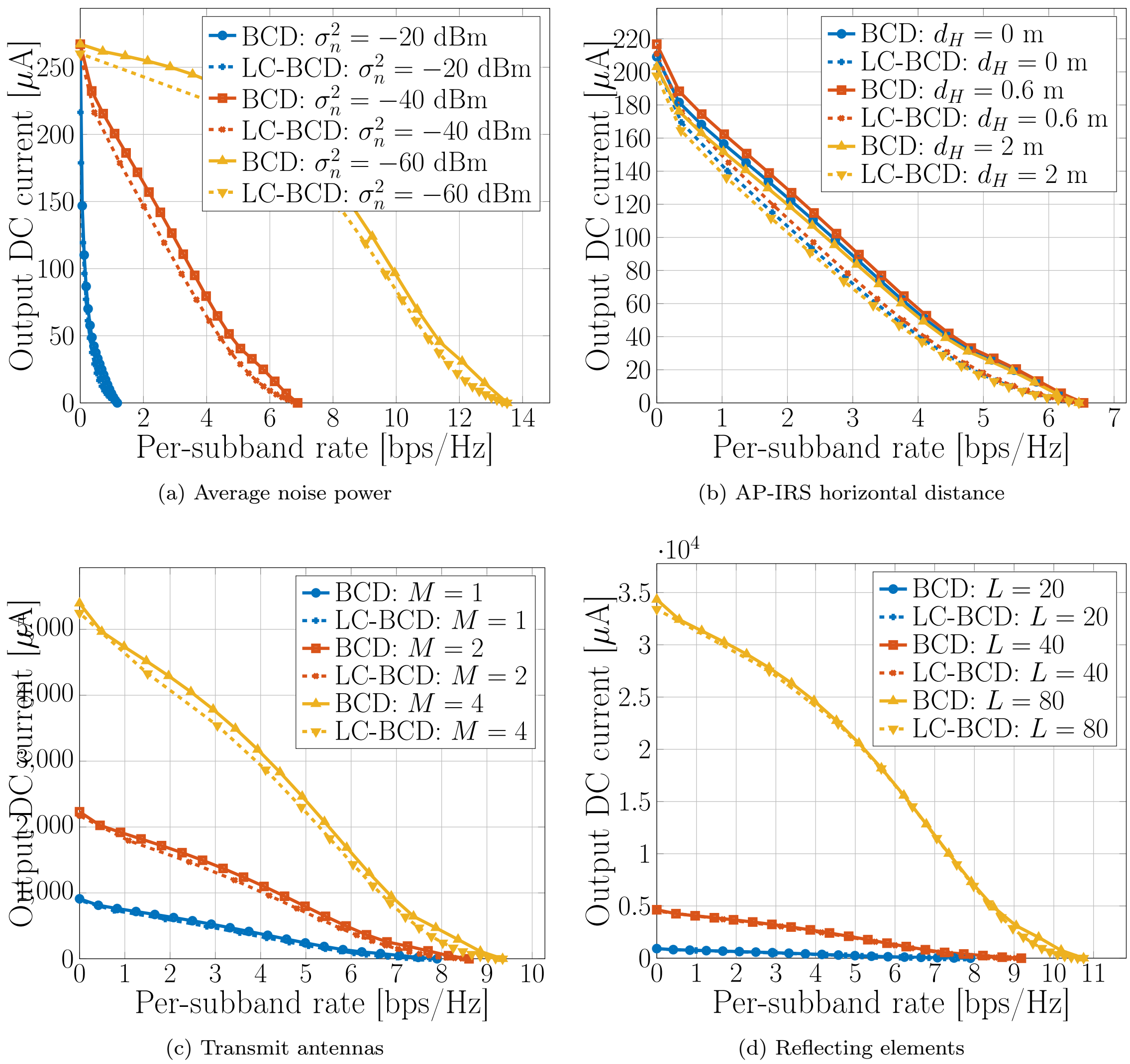
Comment 2.9 — *Minor comment: The size of all the numerical results figures is too small.*

Response: We sincerely apologize for the inconvenience. The figures would be amplified in future versions.

Reviewer 3

Comment 3.1 — *First of all, motivations of studying the IRS on SWIPT is very unclear to me. Please clarify.*

Response: We appreciate your suggestion and have modified the manuscript to emphasize this point. Both SWIPT and IRS aim to achieve spectrum and energy efficient communication, and the effective channel



(a) Average noise power

(b) AP-IRS horizontal distance

(c) Transmit antennas

(d) Reflecting elements

Figure 2: Average achievable R-E region under different configurations.

enhancement and low power consumption of IRS are expected to bring more opportunities to SWIPT. A major challenge for SWIPT is that information decoding and energy harvesting have different sensitivity on the received signal. Although conventional radios as Wi-Fi and Bluetooth can support a signal strength of -90 to -100 dBm, most existing harvesters can only capture signals at -20 to -30 dBm power level. Since the transmit power is strictly constrained by government regulations, it is crucial to increase the end-to-end power transfer efficiency to boost the harvested power and extend the system coverage. Therefore, we believe the passive beamforming ability of the IRS can significantly enhance WPT and SWIPT performance.

Comment 3.2 — *Also, the contributions of this work are rather unclear, and thus, those should be better mentioned.*

Response: We have resummarized the contribution of this paper as follows.

First, we propose a novel IRS-aided SWIPT architecture based on joint waveform, active and passive beamforming design. To elevate the rectifier nonlinearity, we superpose a multi-carrier unmodulated power waveform (deterministic multisine) to a multi-carrier modulated information waveform and evaluate the performance under TS and PS modes. The proposed joint waveform, active and passive beamforming architecture exploits rectifier nonlinearity, beamforming gain and channel selectivity across spatial and frequency domains to enlarge the achievable R-E region. This is the first paper to tackle energy harvester nonlinearity, multi-carrier transmission, and co-localized receiver for IRS-aided SWIPT.

Second, we characterize the achievable R-E region with multiple energy maximization problems with different rate constraints, and solve each R-E boundary point by an Block Coordinate Descent (BCD) algorithm based on the Channel State Information at the Transmitter (CSIT). For active beamforming, we prove that the global optimal active information and power precoders coincide at Maximum-Ratio Transmission (MRT) under rectifier nonlinearity. For passive beamforming, we propose a Successive Convex Approximation (SCA) algorithm that obtains the IRS reflection coefficient by relax-then-extract method. Finally, the superposed waveform is optimized by the Geometric Programming (GP) technique. The IRS phase shift, active precoder and waveform amplitude are updated iteratively until convergence.

Third, we introduce two low-complexity adaptive waveform design in the closed form to avoid the exponential complexity of the GP algorithm. The Water-Filling (WF) strategy for modulated waveform and SMF strategy for multisine waveform are combined under TS and PS setups to facilitate practical SWIPT implementation. To cooperate with the low-complexity waveform schemes, we modify the passive beamforming algorithm and characterize the R-E region by varying the duration ratio for TS and the waveform and splitting ratios for PS. The performance of the low-complexity design is proved close to the BCD algorithm under different configurations.

Fourth, we provide numerical results to evaluate the performance of the proposed algorithms. It is concluded that 1) the proposed BCD algorithm always converge to local optimal points, 2) multisine waveform is beneficial to multi-carrier energy transfer especially when the number of subbands is large, 3) TS is preferred at low Signal-to-Noise Ratio (SNR) while PS is preferred at high SNR, 4) there exist two optimal IRS development locations, one close to the AP and one close to the UE, 5) the output SNR scales linearly with the number of transmit antennas and quadratically with the number of IRS elements, 6) due to rectifier nonlinearity, the output DC current scales quadratically with the number of transmit antennas and quartically with the number of IRS elements, 7) at different R-E points, the optimal IRS reflection coefficients coincide for narrowband SWIPT but differ for broadband SWIPT, 8) the proposed algorithms are robust to practical limitations as inaccurate cascaded CSIT and finite IRS reflection states.

Comment 3.3 — *Please explain the derived results more intuitively for better understanding.*

Response: We have updated the discussion carefully and the reviewer is referred to the updated manuscript.

Comment 3.4 — *Authors assumed the unrealistic situation: the channels are assumed to be perfectly known. However, in practice, the channel should be estimated, e.g., as studied in [R1], [R5]. It would be much better to discuss the channel estimation issue by citing the above references.*

Response: We have added those citations and considered the influence of imperfect cascaded CSIT on the R-E performance. The reviewer is referred to Response 1.1 for details.

Comment 3.5 — *More simulation results should be added to better and aggregately validate the effectiveness of the proposed method.*

Response: Thank you for the suggestion. On top of the previous version, we have introduced two closed-form adaptive waveform design under TS and PS modes and compared the results with the original BCD algorithm. Fig. 2a – 2d demonstrate that the low-complexity BCD algorithm achieves near-optimal performance under different configurations. Moreover, we have investigated the impact of imperfect cascaded CSIT and quantized IRS on the achievable R-E region in Fig. 1 and 3. It can be concluded that the proposed algorithms are robust to CSIT inaccuracy of the AP-IRS-UE link, and even $b = 1$ (i.e. two-state reflection) brings considerable

R-E gain over the benchmark scheme without IRS. These observations demonstrate the advantage of the proposed joint waveform, active and passive beamforming design in practical IRS-aided SWIPT systems.

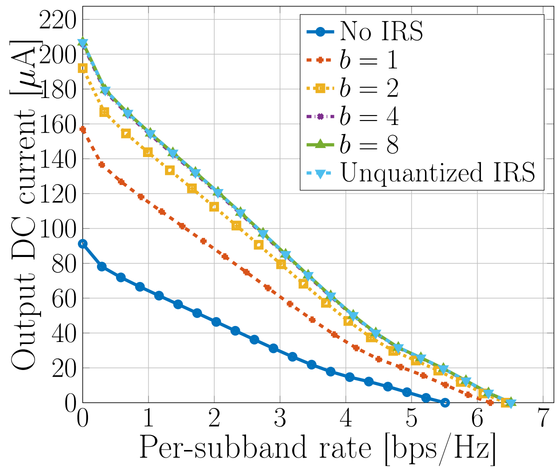


Figure 3: Average R-E region with quantized IRS for $M = 1$, $N = 16$, $L = 20$, $\sigma_n^2 = -40\,\mathrm{dBm}$, $B = 10\,\mathrm{MHz}$ and $d_\mathrm{H} = d_\mathrm{V} = 2\,\mathrm{m}$.

Comment 3.6 — *The sizes of figures are too small.*

Response: We sincerely apologize for the inconvenience. The figures would be amplified in future versions.

References

[R1] J.-M. Kang, "Intelligent Reflecting Surface: Joint Optimal Training Sequence and Refection Pattern," *IEEE Communications Letters*, vol. 24, no. 8, pp. 1784–1788, aug 2020.

[R2] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed Channel Estimation for Intelligent Reflecting Surface-Assisted Millimeter Wave Systems," *IEEE Signal Processing Letters*, vol. 27, pp. 905–909, 2020.

[R3] Q. Wu, X. Zhou, and R. Schober, "IRS-Assisted Wireless Powered NOMA: Do We Really Need Different Phase Shifts in DL and UL?" *arXiv preprint arXiv:2102.08739*, 2021.

[R4] O. Ozdogan, E. Bjornson, and E. G. Larsson, "Intelligent Reflecting Surfaces: Physics, Propagation, and Pathloss Modeling," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 581–585, 2020.

[R5] C. You, B. Zheng, and R. Zhang, "Intelligent Reflecting Surface with Discrete Phase Shifts: Channel Estimation and Passive Beamforming," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*. IEEE, jun 2020, pp. 1–6.