Prof. Liang Xiao
Editor
IEEE Transactions on Communications

Re: Decision on TCOM-TPS-20-1501 (IRS-Aided SWIPT: Joint Waveform, Active and Passive Beamforming Design Under Nonlinear Harvester Model)

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, proved the unit-rank property of the obtained IRS matrix, 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 also 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 link. The proposed passive beamforming algorithms rely on such cascaded CSIT, and recent researches proposed element-wise on/off switching [25], training sequence and reflection pattern design [R1, R2], and compressed sensing [R3] techniques to solve this issue. Those references have been cited in the literature review. To address the comment, we have revised the simulation results and applied the proposed algorithms to an imperfect CSIT model where the estimation of the cascaded link at subband n is

$$\hat{\boldsymbol{V}}_n = \boldsymbol{V}_n + \tilde{\boldsymbol{V}}_n \tag{E1}$$

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. F1 demonstrates the robustness of the proposed passive beamforming algorithm to cascaded CSIT inaccuracy for broadband SWIPT with different number of IRS elements.

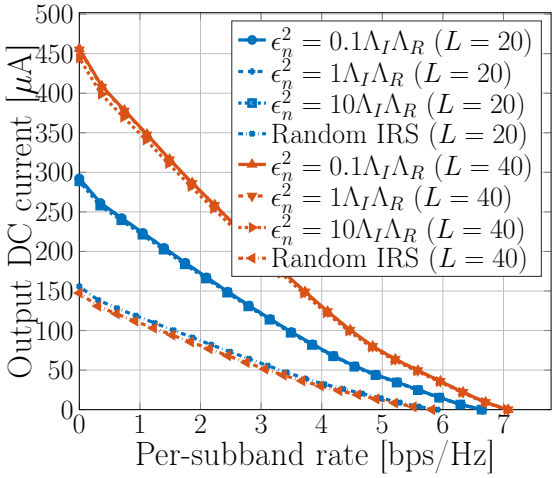


Figure F1: Average R-E region with imperfect cascaded CSIT for $M = 1$, $N = 16$, $\sigma_n^2 = -40\,\text{dBm}$, $B = 10\,\text{MHz}$ and $d_\text{H} = d_\text{V} = 2\,\text{m}$.

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 "don't run before we can walk" 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 be premature 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]. To reflect this aspect, we have revised the title to "IRS-Aided SWIPT: Joint Waveform, Active and Passive Beamforming Design Under Nonlinear Harvester Model".

Third, multi-user WPT leads to its own set of challenges compared with 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, 38]. Consequently, single-user and multi-user WPT are designed differently in [8, 38]. 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. To prove the rank-1 property of the IRS matrix $\mathbf{\Phi}$, we notice that for any feasible $\mathbf{\Phi}$ to problem (26), $\mathrm{tr}(\mathbf{\Phi}) = L + 1$ always holds due to the modulus constraint (26c). Therefore, we add a constant term $-\mathrm{tr}(\mathbf{\Phi})$ to (26a) and recast problem (26) as

$$\begin{aligned}
\max_{\mathbf{\Phi}} \quad & -\mathrm{tr}(\mathbf{\Phi}) + \tilde{z}(\mathbf{\Phi}) && \text{(E2a)}\\
\text{s.t.} \quad & R(\mathbf{\Phi}) \geq \bar{R}, && \text{(E2b)}\\
& \mathrm{diag}^{-1}(\mathbf{\Phi}) = \mathbf{1}, && \text{(E2c)}\\
& \mathbf{\Phi} \succeq \mathbf{0}, && \text{(E2d)}\\
& \mathrm{rank}(\mathbf{\Phi}) = 1. && \text{(E2e)}
\end{aligned}$$

By applying rank constraint relaxation, problem (E2a)–(E2d) is convex with respect to $\mathbf{\Phi}$ and satisfies the Slater's condition [48], thus strong duality holds. The corresponding Lagrangian function at iteration i is

given by

$$\mathcal{L} = \mathrm{tr}\left(\boldsymbol{\Phi}^{(i)}\right) - \frac{1}{2}\beta_2\rho\mathrm{tr}\Big((\boldsymbol{C}_{\mathrm{I},0}+\boldsymbol{C}_{\mathrm{P},0})\boldsymbol{\Phi}^{(i)}\Big) - \frac{3}{8}\beta_4\rho^2\Bigg(4t_{\mathrm{I},0}^{(i-1)}\mathrm{tr}\Big(\boldsymbol{C}_{\mathrm{I},0}\boldsymbol{\Phi}^{(i)}\Big) + \sum_{k=-N+1}^{N-1} 2(t_{\mathrm{P},k}^{(i-1)})^*\mathrm{tr}\Big(\boldsymbol{C}_{\mathrm{P},k}\boldsymbol{\Phi}^{(i)}\Big)$$
$$+ 2(t_{\mathrm{I},0}^{(i-1)} + t_{\mathrm{P},0}^{(i-1)})\mathrm{tr}\Big((\boldsymbol{C}_{\mathrm{I},0}+\boldsymbol{C}_{\mathrm{P},0})\boldsymbol{\Phi}^{(i)}\Big) - \mathrm{tr}^2\Big((\boldsymbol{C}_{\mathrm{I},0}-\boldsymbol{C}_{\mathrm{P},0})\boldsymbol{\Phi}^{(i)}\Big)\Bigg)$$
$$+ \mu\Bigg(2^{\bar{R}} - \prod_{n=1}^{N}\Bigg(1 + \frac{(1-\rho)\mathrm{tr}\Big(\boldsymbol{C}_n\boldsymbol{\Phi}^{(i)}\Big)}{\sigma_n^2}\Bigg)\Bigg) + \mathrm{tr}\Big(\mathrm{diag}(\boldsymbol{\nu})\odot\Big(\boldsymbol{\Phi}^{(i)}\odot\boldsymbol{I}-\boldsymbol{I}\Big)\Big) - \mathrm{tr}\Big(\boldsymbol{\Upsilon}\boldsymbol{\Phi}^{(i)}\Big) + \zeta, \quad \text{(E3)}$$

where μ, $\boldsymbol{\nu}$, $\boldsymbol{\Upsilon}$ denote respectively the scalar, vector and matrix Lagrange multiplier associated with constraint (E2b), (E2c) and (E2d), and ζ collects all terms irrelevant to $\boldsymbol{\Phi}^{(i)}$. The Karush–Kuhn–Tucker (KKT) conditions on the primal and dual solutions are

$$\mu^\star \geq 0, \boldsymbol{\Upsilon}^\star \succeq \mathbf{0}, \quad \text{(E4a)}$$

$$\boldsymbol{\nu}^\star \odot \mathrm{diag}^{-1}(\boldsymbol{\Phi}^\star) = \mathbf{0}, \boldsymbol{\Upsilon}^\star\boldsymbol{\Phi}^\star = \mathbf{0}, \quad \text{(E4b)}$$

$$\nabla_{\boldsymbol{\Phi}^\star}\mathcal{L} = 0. \quad \text{(E4c)}$$

We then derive the gradient explicitly and rewrite (E4c) as

$$\boldsymbol{\Upsilon}^\star = \boldsymbol{I} - \boldsymbol{\Delta}^\star \quad \text{(E5)}$$

with $\boldsymbol{\Delta}^\star$ given by

$$\boldsymbol{\Delta}^\star = \frac{1}{2}\beta_2\rho(\boldsymbol{C}_{\mathrm{I},0}+\boldsymbol{C}_{\mathrm{P},0}) + \frac{3}{4}\beta_4\rho^2\Bigg(2t_{\mathrm{I},0}^{(i-1)}\boldsymbol{C}_{\mathrm{I},0} + \sum_{k=-N+1}^{N-1}(t_{\mathrm{P},k}^{(i-1)})^*\boldsymbol{C}_{\mathrm{P},k} + (t_{\mathrm{I},0}^{(i-1)} + t_{\mathrm{P},0}^{(i-1)})(\boldsymbol{C}_{\mathrm{I},0}+\boldsymbol{C}_{\mathrm{P},0})$$
$$- (\boldsymbol{C}_{\mathrm{I},0}-\boldsymbol{C}_{\mathrm{P},0})\mathrm{tr}\Big((\boldsymbol{C}_{\mathrm{I},0}+\boldsymbol{C}_{\mathrm{P},0})\boldsymbol{\Phi}^\star\Big)\Bigg) + \mu^\star\sum_{n=1}^{N}\frac{(1-\rho)\boldsymbol{C}_n}{\sigma_n^2}\prod_{n'=1,n'\neq n}^{N}\Bigg(1 + \frac{(1-\rho)\mathrm{tr}\Big(\boldsymbol{C}_{n'}\boldsymbol{\Phi}^\star\Big)}{\sigma_{n'}^2}\Bigg)$$
$$- \mathrm{diag}(\boldsymbol{\nu}^\star). \quad \text{(E6)}$$

Note that (E4b) suggests $\mathrm{rank}(\boldsymbol{\Upsilon}^\star) + \mathrm{rank}(\boldsymbol{\Phi}^\star) \leq L+1$. By reusing the proof in Appendix A of [R4], we conclude $\mathrm{rank}(\boldsymbol{\Upsilon}^\star) \geq L$. On the other hand, $\boldsymbol{\Phi}^\star$ cannot be zero matrix and $\mathrm{rank}(\boldsymbol{\Phi}^\star) \geq 1$. Therefore, any optimal solution $\boldsymbol{\Phi}^\star$ to the relaxed problem (E2) satisfies $\mathrm{rank}(\boldsymbol{\Phi}^\star) = 1$. Due to the equivalence between (26a) and (E2a), $\boldsymbol{\Phi}^\star$ is also optimal to the relaxed problem (26). We have clarified this point in Proposition 1.

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 have been amplified in the revised manuscript.

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 non-convex unit-rank constraint, if the proof to Proposition 1 is unavailable, existing approaches include 1) relax the constraint and use Gaussian randomization method to obtain a high-quality solution [39], 2) replace the unit-rank constraint with the constraint on largest singular value $\mathrm{Tr}(\boldsymbol{\Phi}) - \sigma_{\max}(\boldsymbol{\Phi}) \leq 0$ then solve the problem by the penalty method [33], [R5], 3) apply the gradient descent in the manifold space [R6]. All three 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. We have mentioned in the conclusion that "Also, the use of more involved IRS architecture based on partially/fully connected and/or direction- and frequency-dependent reflection could be considered in future works."

Comment 2.2 — *Why is MRT considered as precoder by* (28) *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

$$\boldsymbol{w}_{\mathrm{I/P},n} = s_{\mathrm{I/P},n}\boldsymbol{b}_{\mathrm{I/P},n} \tag{E7}$$

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

$$\boldsymbol{b}^{\star}_{\mathrm{I/P},n} = \frac{\boldsymbol{h}_n}{\|\boldsymbol{h}_n\|} \tag{E8}$$

From the perspective of WIT, the MRT precoder maximizes $|\boldsymbol{h}_n^H \boldsymbol{w}_{\mathrm{I},n}| = \|\boldsymbol{h}_n\| s_{\mathrm{I},n}$ thus maximizes the rate (8). From the perspective of WPT, the MRT precoder maximizes $(\boldsymbol{h}_n^H \boldsymbol{w}_{\mathrm{I/P},n})(\boldsymbol{h}_n^H \boldsymbol{w}_{\mathrm{I/P},n})^* = \|\boldsymbol{h}_n\|^2 s^2_{\mathrm{I/P},n}$ 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, summarized the conclusion in Proposition 3, and add the proof above 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 link on the R-E performance, and Fig. F1 shows that the proposed algorithms are robust to CSIT inaccuracy for different L.

Comment 2.4 — *All the assumptions and relaxations adopted used in the derivation of results as in* (25) *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. (26b)–(26e)). Also, (24) holds due to symmetry. Hence, the quadratic objective function (25) is obtained by plugging (21)–(24) into (20). Solving relaxed problem (26) iteratively is guaranteed to converge to a local optimal point of the original passive beamforming problem, and the proof is provided in Appendix B.

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 provide local optimal solutions with local convergence proof, and the performance indeed depends on the initialization. For the passive beamforming Algorithm 1 and 3, we initialize the phase shift of all IRS elements as i.i.d. uniform random variables over $[0, 2\pi)$. For Algorithm 2, we initialize the modulated waveform by the Water-Filling (WF) strategy E9 and the multisine waveform by the Scaled Matched Filter (SMF) scheme E10, and assume $\rho^{(0)} = \bar{\rho}^{(0)} = 1$. These parameters are used for general initialization and regulated automatically afterwards. However, as Algorithm 1, 2 and 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 4 and 5 by the solution at the previous point, then perform the optimization again. For a tolerance of $\epsilon = 10^{-8}$, Algorithm 1–5 typically converge within 2, 7, 3, 2, 3 iterations, respectively.

We also would like to point out that, even in the absence of IRS, the problem of joint waveform and beamforming design for SWIPT is also non-convex and algorithms only converging toward a local optimum have been identified, with the global optimum remaining unsolved.

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 (26) is not a Semidefinite Programming (SDP) due to the squared term in the objective function. Hence, existing complexity analysis tools cannot be direct applied to Algorithm 1 and 3. For Algorithm 2, the computational complexity scales exponentially with the number of subbands [38, 42] but an analytical expression cannot be derived on top of relevant literatures. To facilitate practical SWIPT implementation and address the reviewer's comment, we have added in the revised manuscript two closed-form adaptive waveform amplitude schemes by combining WF and SMF in the time and power domains. For WIT, the optimal WF strategy assigns the amplitude of modulated tone n by

$$s_{\mathrm{I},n} = \sqrt{2\left(\lambda - \frac{\sigma_n^2}{P\|\boldsymbol{h}_n\|^2}\right)^+} \tag{E9}$$

where λ 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 [43], 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

$$s_{\mathrm{P},n} = \sqrt{\frac{2P}{\sum_{n=1}^{N}\|\boldsymbol{h}_n\|^{2\alpha}}}\|\boldsymbol{h}_n\|^{\alpha} \tag{E10}$$

where the scaling ratio $\alpha \geq 1$ is given and can be adjusted to exploit the rectifier nonlinearity and frequency selectivity. When the receiver works in TS mode, there is no superposition in the suboptimal waveform design. Modulated waveform with amplitude (E9) is used in the data session while multisine waveform with amplitude (E10) is used in the energy session. When the receiver works in PS mode, we jointly design the combining ratio δ with the splitting ratio ρ, and assign the superposed waveform amplitudes as

$$s_{\mathrm{I},n} = \sqrt{2(1-\delta)\left(\lambda - \frac{\sigma_n^2}{P\|\boldsymbol{h}_n\|^2}\right)^{+}}, \tag{E11}$$

$$s_{\mathrm{P},n} = \sqrt{\frac{2\delta P}{\sum_{n=1}^{N}\|\boldsymbol{h}_n\|^{2\alpha}}}\|\boldsymbol{h}_n\|^{\alpha}. \tag{E12}$$

Besides, minor modifications are required for passive beamforming to accommodate both low-complexity waveform schemes. To achieve the WIT point ($\eta = \rho = 0$), the rate (19) instead of the DC current (25) should be maximized. To achieve any non-WIT point (i.e., $z \neq 0$), the rate constraint (26b) should be dropped as the R-E region is instead characterized by varying either η or δ and ρ. The Modified-SCA (M-SCA) Algorithm A1 summarizes the modified passive beamforming design when the receiver works in PS mode. The tightness of rank relaxation and local optimality are proved similar to Proposition 1 and 2 thus omitted here. For the low-complexity design, when the receiver works in PS mode, we instead obtain the phase shift

Algorithm A1 M-SCA: IRS Phase Shift.

Input $\beta_2, \beta_4, \boldsymbol{h}_{\mathrm{D},n}, \boldsymbol{V}_n, \sigma_n, \boldsymbol{w}_{\mathrm{I/P},n}, \rho, \epsilon, \forall n$
Construct $\boldsymbol{V}, \boldsymbol{M}, \boldsymbol{M}_n, \boldsymbol{C}_n, \boldsymbol{C}_{\mathrm{I/P},k}, \forall n, k$
Initialize $i \leftarrow 0$, $\boldsymbol{\Phi}^{(0)}$
if $\rho = 0$ **then**
    Get $\boldsymbol{\Phi}^{\star}$ by maximizing (19) s.t. (26c), (26d)
else
    Set $t_{\mathrm{I/P},k}^{(0)}, \forall k$ by (18)
    repeat
        $i \leftarrow i + 1$
        Get $\boldsymbol{\Phi}^{(i)}$ by maximizing (25) s.t. (26c), (26d)
        Update $t_{\mathrm{I/P},k}^{(i)}, \forall k$ by (18)
        Compute $z^{(i)}$ by (20)
    until $|z^{(i)} - z^{(i-1)}| \leq \epsilon$
    Set $\boldsymbol{\Phi}^{\star} = \boldsymbol{\Phi}^{(i)}$
end if
Get $\bar{\boldsymbol{\phi}}^{\star}$ by eigen decomposition, $\boldsymbol{\Phi}^{\star} = \bar{\boldsymbol{\phi}}^{\star}(\bar{\boldsymbol{\phi}}^{\star})^H$
Set $\theta_l^{\star} = \arg(\bar{\phi}_l^{\star}/\bar{\phi}_{L+1}^{\star}), \forall l$, $\boldsymbol{\phi}^{\star} = [e^{j\theta_1^{\star}}, \ldots, e^{j\theta_L^{\star}}]^H$
Output $\boldsymbol{\phi}^{\star}$

by Algorithm A1 and the waveform amplitude by (E11) and (E12). The steps are summarized in the Low Complexity-BCD (LC-BCD) Algorithm A2. Fig. F2a–F2d demonstrate that the LC-BCD Algorithm A2 achieves near-optimal performance under different configurations.

Algorithm A2 LC-BCD: Waveform and Beamforming.

Input $\beta_2, \beta_4, \boldsymbol{h}_{\mathrm{D},n}, \boldsymbol{V}_n, P, \sigma_n, \delta, \rho, \epsilon, \forall n$
Initialize $i \leftarrow 0$, $\boldsymbol{\phi}^{(0)}, \boldsymbol{b}_{\mathrm{I/P},n}^{(0)}, \boldsymbol{s}_{\mathrm{I/P}}^{(0)}, \forall n$
Set $\boldsymbol{w}_{\mathrm{I/P},n}^{(0)}, \forall n$ by (27)
repeat
    $i \leftarrow i + 1$
    Get $\boldsymbol{\phi}^{(i)}$ based on $\boldsymbol{w}_{\mathrm{I/P}}^{(i-1)}$ by Algorithm A1
    Update $\boldsymbol{h}_n^{(i)}, \boldsymbol{b}_n^{(i)}, \forall n$ by (5), (28)
    Update $\boldsymbol{s}_{\mathrm{I}}^{(i)}, \boldsymbol{s}_{\mathrm{P}}^{(i)}$ by (E11), (E12)
    Update $\boldsymbol{w}_{\mathrm{I/P},n}^{(i)}, \forall n$ by (27)
    Compute $R^{(i)}, z^{(i)}$ by (29), (30)
until $|R^{(i)} - R^{(i-1)}| \leq \epsilon$ or $|z^{(i)} - z^{(i-1)}| \leq \epsilon$
Set $\boldsymbol{\phi}^\star = \boldsymbol{\phi}^{(i)}, \boldsymbol{w}_{\mathrm{I/P}}^\star = \boldsymbol{w}_{\mathrm{I/P}}^{(i)}$
Output $\boldsymbol{\phi}^\star, \boldsymbol{w}_{\mathrm{I}}^\star, \boldsymbol{w}_{\mathrm{P}}^\star$

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 ideal but common in microwave theory (e.g., [R7]). Based on the circuit analysis, the authors of [24] shows that the reflection amplitude depends on the phase shift, which is minimized at 0 phase shift and is negatively correlated to the element resistance. For $R_l = 1\,\Omega$, $|\phi_l|$ is 1 for $\theta_l = \pi$ but around 0.6 for $\theta_l = 0$. From the perspective of field and propagation, [R8] 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 lossless reflection pattern at the IRS in this paper.

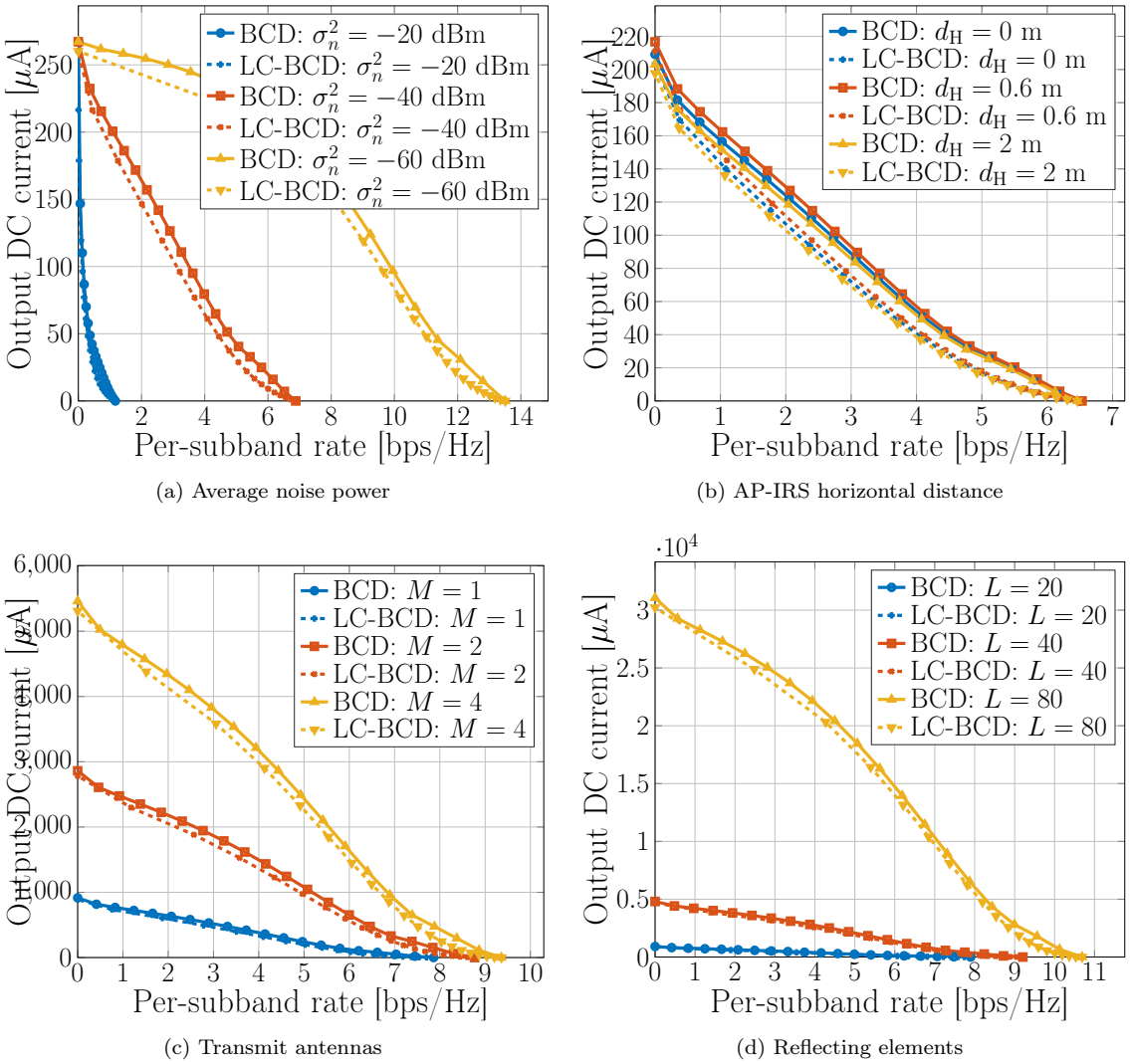
Comment 2.9 — *Minor comment: The size of all the numerical results figures is too small.*

Response: We sincerely apologize for the inconvenience. The figures have been amplified in the revised manuscript.

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. The major issue of WPT and SWIPT is the low harvested DC power. To increase it, boosting the harvester input power level would be very helpful. This is achieved by the help of IRS. Indeed, both SWIPT and IRS aim to achieve spectrum and energy efficient communication, and the smart channel control 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 $-100\,\mathrm{dBm}$, most existing harvesters can only capture signals at -20 to $-30\,\mathrm{dBm}$ 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.



(a) Average noise power

(b) AP-IRS horizontal distance

(c) Transmit antennas

(d) Reflecting elements

Figure F2: Average achievable R-E region under different configurations.

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 make an efficient use of 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 receiving modes. The proposed joint waveform, active and passive beamforming architecture exploits the rectifier nonlinearity, a beamforming gain, and the channel selectivity across spatial and frequency domains to enlarge the achievable R-E region. This is the first paper to propose a joint waveform, active and passive beamforming architecture for IRS-aided SWIPT. By doing so, it is also the first paper to properly model the harvester nonlinearity (including its impact on both power and shape of the incoming waveform) and account for its crucial role in IRS-aided SWIPT.

Second, we characterize each R-E boundary point by energy maximization under rate constraint, and solve the problem by a 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) even with rectifier nonlinearity. For passive beamforming, we propose a Successive Convex Approximation (SCA) algorithm and retrieve the IRS phase shift by eigen decomposition with optimality proof. 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. This is the first paper to jointly optimize waveform and active/passive beamforming in IRS-aided SWIPT.

Third, we introduce two closed-form adaptive waveform schemes to avoid the exponential complexity of the GP algorithm. The Water-Filling (WF) strategy for modulated waveform and the SMF strategy for multisine waveform are combined in the time domain and power domain to facilitate practical SWIPT implementation. To accommodate the suboptimal waveform schemes, we modify the passive beamforming algorithm and characterize the R-E region by varying the duration ratio under TS mode and the combining and splitting ratios under PS mode. The proposed low-complexity designs achieve near-optimal R-E performance in different scenarios.

Fourth, we provide numerical results to evaluate the proposed algorithms. It is concluded that 1) multisine waveform is beneficial to energy transfer especially when the number of subbands is large, 2) TS is preferred at low Signal-to-Noise Ratio (SNR) while PS is preferred at high SNR, 3) there exist two optimal IRS development locations, one close to the AP and one close to the UE, 4) the output SNR scales linearly with the number of transmit antennas and quadratically with the number of IRS elements, 5) due to the rectifier nonlinearity, the output DC current scales quadratically with the number of transmit antennas and quartically with the number of IRS elements, 6) for narrowband SWIPT, the optimal active and passive beamforming for any R-E point are also optimal for the whole R-E region, 7) for broadband SWIPT, the optimal active and passive beamforming depend on specific R-E point and require adaptive designs, 8) the proposed algorithms are robust to practical impairments such 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, R2]. It would be much better to discuss the channel estimation issue by citing the above references.*

Response: We have added those references and considered the influence of imperfect cascaded CSIT on the R-E performance. The reviewer is referred to Response 1.1 for details. Simulation result in Fig. F1 demonstrates the robustness of the proposed joint waveform, active and passive beamforming algorithm to cascaded CSIT inaccuracy with different number of IRS elements.

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. To address this comment, we have introduced two closed-form adaptive waveform design and compared the results with the original BCD Algorithm 4. Fig. F2a–F2d demonstrate that the LC-BCD Algorithm A2 achieves near-optimal performance under different configurations. Moreover, we have investigated in Fig. F1 and F3 the impact of imperfect cascaded CSIT and quantized IRS on the achievable R-E region. It can be concluded that the proposed algorithms are robust to CSIT inaccuracy of the cascaded 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.

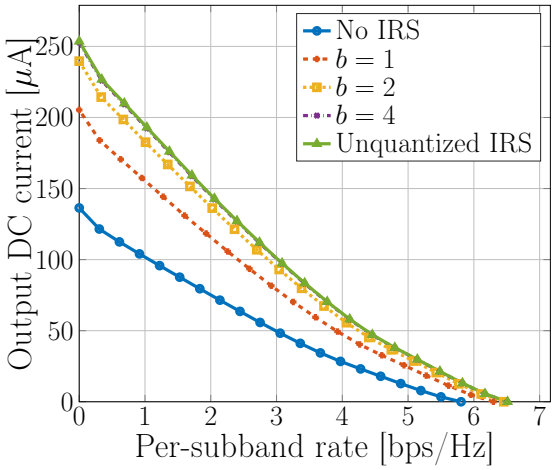


Figure F3: Average R-E region with quantized IRS for $M = 1$, $N = 16$, $L = 20$, $\sigma_n^2 = -40\,\text{dBm}$, $B = 10\,\text{MHz}$ and $d_\text{H} = d_\text{V} = 2\,\text{m}$.

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

Response: We sincerely apologize for the inconvenience. The figures have been amplified in the revised manuscript.

References

[R1] 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.

[R2] 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.

[R3] 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.

[R4] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, "Resource allocation for irs-assisted full-duplex cognitive radio systems," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7376–7394, Dec. 2020.

[R5] 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.

[R6] J. Zuo, Y. Liu, L. Yang, L. Song, and Y.-C. Liang, "Reconfigurable intelligent surface enhanced noma assisted backscatter communication system," *arXiv preprint arXiv:2012.10111*, pp. 1–6, Dec. 2020.

[R7] S. Shen, B. Clerckx, and R. Murch, "Modeling and architecture design of intelligent reflecting surfaces using scattering parameter network analysis," *arXiv preprint arXiv:2011.11362*, pp. 1–30, Nov. 2020.

[R8] 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, May 2020.