Prof. Liang Xiao
Editor
IEEE Transactions on Communications

Re: Decision on TCOM-TPS-21-0553 (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 added computational complexity analysis for all proposed algorithms, addressed the optimality issue regarding Algorithm 4, discussed the convergence conditions in Algorithm 5, emphasized the design challenge introduced by passive beamforming and nonlinear EH model, and updated the simulation parameters reference list. To answer the questions and concerns, we also include a point-to-point response below to discuss 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

Editor

Comment 0.1 — *Please double-check Proposition 5 and the algorithm may not converge to a local optimal due to the coupling constraint* (15b).

Response: Thank you for pointing this out. The editor is referred to Response 3.1.

Comment 0.2 — *Please discuss the convergence criteria in details and justify the formulated optimization problem.*

Response: Please refer to Response 1.4.

Comment 0.3 — *How does the IRS passive beamforming impact on the waveform design compared with existing systems?*

Response: Please refer to Response 2.2.

Comment 0.4 — *In the introduction, it is suggested to also discuss the new design challenges posed by the nonlinear model compared to the existing designs under the linear model.*

Response: Please refer to Response 2.3.

Reviewer 1

Comment 1.1 — *As this paper proposed a BCD-based algorithm and LC-BCD algorithm, it could be better if the authors can further discuss the complexity and provide some measurements or some simulation results.*

Response: We appreciate your suggestion and agree that complexity analysis is necessary in algorithm evaluation.

For SCA Algorithm 1, in the previous manuscript, we applied SCA on $t_{\mathrm{I},0}t_{\mathrm{P},0}$ at iteration i as

$$\begin{aligned} t_{\mathrm{I},0}^{(i)}t_{\mathrm{P},0}^{(i)} &= \frac{1}{4}(t_{\mathrm{I},0}^{(i)}+t_{\mathrm{P},0}^{(i)})^2-\frac{1}{4}(t_{\mathrm{I},0}^{(i)}-t_{\mathrm{P},0}^{(i)})^2 \\ &\geq \frac{1}{2}(t_{\mathrm{I},0}^{(i)}+t_{\mathrm{P},0}^{(i)})(t_{\mathrm{I},0}^{(i-1)}+t_{\mathrm{P},0}^{(i-1)})-\frac{1}{4}(t_{\mathrm{I},0}^{(i-1)}+t_{\mathrm{P},0}^{(i-1)})^2-\frac{1}{4}(t_{\mathrm{I},0}^{(i)}-t_{\mathrm{P},0}^{(i)})^2, \end{aligned} \tag{E1}$$

which is generally feasible and widely adopted when $t_{\mathrm{I},0}$ and $t_{\mathrm{P},0}$ are variables. Hence, we claimed Algorithm 1 "is not a Semidefinite Programming (SDP) due to the quadratic objective function (24)". During revision, we realized that $t_{\mathrm{I/P},0}=\mathrm{tr}(\boldsymbol{C}_{\mathrm{I/P},0}\boldsymbol{\Phi})$ are essentially linear real expressions related to optimization variable $\boldsymbol{\Phi}$, and the SCA can be further reduced to

$$t_{\mathrm{I},0}^{(i)}t_{\mathrm{P},0}^{(i)}\geq t_{\mathrm{I},0}^{(i)}t_{\mathrm{P},0}^{(i-1)}+t_{\mathrm{P},0}^{(i)}t_{\mathrm{I},0}^{(i-1)}-t_{\mathrm{I},0}^{(i-1)}t_{\mathrm{P},0}^{(i-1)}. \tag{E2}$$

Therefore, the quadratic term in the objective function (24) is eliminated and the updated problem 25 is indeed an SDP. Given a solution accuracy ϵ_{IPM} for the interior-point method, the computational complexity of Algorithm 1 is $\mathcal{O}\left(I_{\mathrm{SCA}}(L+2)^4(L+1)^{0.5}\log(\epsilon_{\mathrm{IPM}}^{-1})\right)$, where I_{SCA} denotes the number of SCA iterations [43]. We have revised the manuscript as follows.

The passive beamforming design is summarized in the SCA Algorithm 1, where the relaxed problem (25a)–(25d) involves a $(L+1)$-order positive semi-definite matrix variable and $(L+2)$ linear constraints. Given a solution accuracy ϵ_{IPM} for the interior-point method, the computational complexity of Algorithm 1 is $\mathcal{O}\left(I_{\mathrm{SCA}}(L+2)^4(L+1)^{0.5}\log(\epsilon_{\mathrm{IPM}}^{-1})\right)$, where I_{SCA} denotes the number of SCA iterations [43].

For GP Algorithm 2, the computational complexity is exponential w.r.t. the number of subbands [46].

The joint waveform amplitude and splitting ratio design is summarized in the GP Algorithm 2, which achieves local optimality at the cost of exponential computational complexity [46].

For M-SCA Algorithm 3, each optimization problem involved is an SDP with $(L+1)$ variables and linear constraints, and the computational complexity is $\mathcal{O}\left(I_{\mathrm{M-SCA}}(L+1)^{4.5}\log(\epsilon_{\mathrm{IPM}}^{-1})\right)$ where $I_{\mathrm{M-SCA}}$ denotes the number of M-SCA iterations.

Compared with Algorithm 1, the rate constraint (25b) is dropped and each SDP in Algorithm 3 involves $(L+1)$ linear constraints. Given a solution accuracy ϵ_{IPM} for the interior-point method, the computational complexity of Algorithm 3 is $\mathcal{O}\left(I_{\mathrm{M-SCA}}(L+1)^{4.5}\log(\epsilon_{\mathrm{IPM}}^{-1})\right)$, where $I_{\mathrm{M-SCA}}$ denotes the number of M-SCA iterations [43]. Note that $I_{\mathrm{M-SCA}}=1$ for the WIT point where no SCA is involved.

For BCD Algorithm 4, since Algorithm 2 is involved per iteration, the computational complexity is exponential w.r.t. the number of subbands.

The steps are summarized in the BCD Algorithm 4, whose computational complexity is exponential as inherited from Algorithm 2.

For LC-BCD Algorithm 5, both waveform and active beamforming are obtained in closed form, and the computational complexity of Algorithm 5 is $\mathcal{O}\left(I_{\mathrm{LC-BCD}}I_{\mathrm{M-SCA}}(L+1)^{4.5}\log(\epsilon_{\mathrm{IPM}}^{-1})\right)$, where $I_{\mathrm{LC-BCD}}$ denotes the number of LC-BCD iterations [43].

Given a solution accuracy ϵ_{IPM} for the interior-point method, the computational complexity of Algorithm 5 is $\mathcal{O}\left(I_{\mathrm{LC-BCD}}I_{\mathrm{M-SCA}}(L+1)^{4.5}\log(\epsilon_{\mathrm{IPM}}^{-1})\right)$, where $I_{\mathrm{LC-BCD}}$ denotes the number of LC-BCD iterations [43].

Comment 1.2 — *Since this paper focused on IRS-aided SWIPT systems, it could be better if the authors can further enrich the survey part to make the paper more comprehensive. In particular, some papers have considered this kind of problem from a different aspect such as [R1]. It is kindly suggested to briefly discuss the main difference between [R1] and the problem considered in this paper.*

Response: Thank you for sharing this paper. Instead of optimizing the phase shift of each IRS element, it tailored a tile-based two-stage optimization framework for large-scale IRS to enable a flexible performance-complexity tradeoff. The model also adapts well to practical IRS models accounting AoA and AoD. In contrast, our paper is more focused on waveform design and rate-energy tradeoff for one SWIPT user with practical co-located information decoder and energy harvester. We have updated the manuscript as follows.

In [R1], the authors proposed a scalable resource allocation framework for SWIPT systems involving large-scale IRS, where the reflectors are grouped into tiles and the optimization process is divided into an offline mode-design stage and an online mode-selection stage. It was concluded that the proposed two-stage algorithm not only enables a flexible balance between performance and complexity but also adapts well to practical IRS and harvester models.

... To the best of our knowledge, all existing papers considered resource allocation and beamforming design for dedicated information and energy users in a single-carrier network. In this paper, we instead build our design based on a proper nonlinear harvester modeling that captures the dependency of the output DC power on both the power and shape of the input waveform, and marry the benefits of joint multi-carrier waveform and active beamforming optimization for SWIPT with the passive beamforming capability of IRS, to investigate the R-E tradeoff for one SWIPT user with practical co-located information decoder and energy harvester.

Comment 1.3 — *It could be better if the authors can provide some references for the parameter values chosen in the simulation part.*

Response: The system layout follows [23]. For the center frequency, we consider 2.4 GHz due to its ubiquitous and license-free properties. Note that the path loss can be greatly reduced (and the R-E performance can be improved) if the system operates at much lower frequency, as considered in most IRS papers. In the previous manuscript, we made a mistake in the choice of reference path loss. As indicated by Comment 3.3, even the free-space path loss should be around 40 dB at 1 m. To fix this issue, we updated the path loss model as IEEE TGn channel model D that is representative of small environments such as residential homes and small offices [48]. It has three clusters with 50 ns r.m.s. delay spread, and the path loss exponent is set to 2 (i.e., free-space model) up to breakpoint distance 10 m and set to 3.5 onwards. This has a similar effect to most IRS literatures that assign a relatively large path loss exponent to the direct path and smaller ones to the incident and reflected paths to penalize the channel with large distance. We also raise the EIRP from 36 dBm to 40 dBm (which still complies with FCC regulations [R2]) and change the receive antenna gain from 2 dBi to 3 dBi. As for the rectenna parameters, the parameters are taken from [8] and the reference have been added to the revised manuscript.

Comment 1.4 — *The reviewer noticed that in Algorithm 5, there are two convergence criteria. It could be better if the authors can briefly interpret this issue.*

Response: We incorporated two cases (WIT/non-WIT) into Algorithm 5, where Algorithm 3 is called during each iteration. To obtain the WIT point, the rate (18) instead of the DC current (24) should be maximized, such that the current expression is dropped and no SCA is involved. To clarify this point, we have updated the description on Algorithm 3 as follows.

To achieve the WIT point, the rate (18) instead of the DC current (24) should be maximized. In such case, the current expression is dropped and no SCA is involved.

Correspondingly, to achieve the WIT point, the LC-BCD Algorithm 5 should maximizes R instead of z. Line 12–17 was modified to avoid ambiguity and Algorithm 5 is indexed below by A1 . The description on Algorithm 5 have been updated as follows.

Algorithm A1 LC-BCD: Waveform and Beamforming.

Input β_2, β_4, $\boldsymbol{h}_{\mathrm{D},n}$, $\boldsymbol{V}_n$, P, σ_n, δ, ρ, ϵ, $\forall n$
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 (26)
Compute $R^{(0)}$, $z^{(0)}$ by (28), (29)
repeat
    $i \leftarrow i + 1$
    Get $\boldsymbol{\phi}^{(i)}$ based on $\boldsymbol{w}_{\mathrm{I/P}}^{(i-1)}$ by Algorithm 3
    Update $\boldsymbol{h}_n^{(i)}$, $\boldsymbol{b}_n^{(i)}$, $\forall n$ by (5), (27)
    Update $\boldsymbol{s}_{\mathrm{I}}^{(i)}$, $\boldsymbol{s}_{\mathrm{P}}^{(i)}$ by (39), (40)
    Update $\boldsymbol{w}_{\mathrm{I/P},n}^{(i)}$, $\forall n$ by (26)
    Compute $R^{(i)}$, $z^{(i)}$ by (28), (29)
    if $\rho = 0$ **then**
        $\Delta = R^{(i)} - R^{(i-1)}$
    else
        $\Delta = z^{(i)} - z^{(i-1)}$
    end if
until $|\Delta| \leq \epsilon$
Set $\boldsymbol{\phi}^{\star} = \boldsymbol{\phi}^{(i)}$, $\boldsymbol{w}_{\mathrm{I/P}}^{\star} = \boldsymbol{w}_{\mathrm{I/P}}^{(i)}$
Output $\boldsymbol{\phi}^{\star}$, $\boldsymbol{w}_{\mathrm{I}}^{\star}$, $\boldsymbol{w}_{\mathrm{P}}^{\star}$

Note that the BCD algorithm obtains the R-E region by varying the rate constraint from 0 to maximum capacity $C_{\max}$[1], and the LC-BCD algorithm characterizes the R-E tradeoff via two-dimensional search over δ, ρ from 0 to 1. For the WIT point, $C_{\max}$ can be obtained in the special case where $\rho = 0$ and the objective function becomes R instead of z.

Reviewer 2

Comment 2.1 — *Although this paper reveals many useful insights, most of them are obtained via simulation and then explained via text. The overall theoretical contributions can be greatly improved if the authors can analytically verify or prove some of them, especially those unique to this paper. The authors may conduct the analysis under some simplified or special cases. If some of the observations and findings have been verified in existing papers, the authors can cite and echo these papers wherever necessary.*

Response: We believe the reviewer was referring to the array gain and scaling law suggested by Figs. 10b and 11b. Recall that M refers to the number of transmit antenna, N refers to the number of subbands, and L refers to the number of IRS elements. From the perspective of WIT, coherent transmission achieves an array gain of M [47] and coherent reflection achieves an array gain of L^2 [23]. From the perspective of WPT, [8] and [14] analytically proved the scaling law of the modulated and multisine waveforms. Even if a suboptimal waveform scheme (uniform power allocation in the frequency domain and matched filter in the spatial domain) is used, when M is sufficiently large, the expectation of the DC current scales quadratically with M. We have cited those papers in the revised manuscript.

[1]Recall in Remark 2 that passive beamforming enables a resource allocation opportunity at the channel such that different capacities are achievable.

Next, we verify the last observation where the expectation of the DC current scales quartically with L in a much simplified scenario. Consider a SISO WPT system in a frequency-flat channel. We assume a sufficiently large L such that the auxiliary link is dominant and the contribution of direct channel is omitted. With uniform power allocation to multisine ($s_{\mathrm{P},n} = \sqrt{2P}/\sqrt{N}$) in the frequency domain, the DC current writes as

$$z_{\mathrm{UP}} = \beta_2(\Lambda_{\mathrm{I}}\Lambda_{\mathrm{R}}L|X|)^2P + \frac{3}{8}\beta_4(\Lambda_{\mathrm{I}}\Lambda_{\mathrm{R}}L|X|)^4F, \tag{E3}$$

where β_2 and β_4 are fixed rectenna parameters, Λ_{I} and Λ_{R} are the path loss of incident and reflected links, X is the random variable that models the small-scale double fading through one IRS element at any subband, and

$$F = \sum_{\substack{n_1,n_2,n_3,n_4 \\ n_1+n_2=n_3+n_4}} s_{n_1}s_{n_2}s_{n_3}s_{n_4} = \frac{4(2N^2+1)}{3N}P^2. \tag{E4}$$

Taking expectation over $|X|$, we have

$$\bar{z}_{\mathrm{UP}} = \beta_2\Lambda_{\mathrm{I}}^2\Lambda_{\mathrm{R}}^2L^2\mathbb{E}\{|X|^2\}P + \frac{3}{8}\beta_4\Lambda_{\mathrm{I}}^4\Lambda_{\mathrm{R}}^4L^4\mathbb{E}\{|X|^4\}\frac{2N^2+1}{2N}P^2, \tag{E5}$$

which verifies the scaling order of L^4. The analysis is not presented in the manuscript due to the space constraint.

Comment 2.2 — *The authors are suggested to study the impact of IRS passive beamforming on the waveform design without IRS. For example, they can show the optimized waveforms with versus without IRS to unveil the interplay between them. Or even better, analytically compare the two waveforms.*

Response: Thank you for the suggestion. Too see this, we first show one realization of direct channel amplitude (without IRS) and composite channel amplitude (with IRS) for WIT and WPT in Fig. F1. It is observed that under the configuration of choice, the WPT-optimized passive beamforming tends to concentrate the channel energy to relatively strong subbands, while the WIT-optimized passive beamforming achieves some balance among all subbands (similar to water-filling at high SNR).

On the other hand, the corresponding amplitude of modulated and multisine waveforms are shown in Fig. F2. Note that no multisine is used for WIT and the case is not shown. It is observed that although passive beamforming can enable a flexible resource allocation opportunity at the channel (please refer to Response 2.10 for details), its impact on waveform design is unobvious.

We have modified the manuscript as follows to emphasize this point.

> Fig. 7a reveals how passive beamforming influence the channel amplitude and waveform design for one channel realization. *First,* the passive beamforming enables a resource allocation opportunity at the channel as discussed in Remark 2. Under the specified configuration, the WPT-optimized IRS concentrates channel energy on strong subbands to widen the strength disparity. Correspondingly, the transmit power is focused on strong subbands to improve the energy efficiency, as observed in 7c and 7d. On the other hand, the WIT-optimized IRS provides a fair gain over all subchannels when L is sufficiently large. This is reminiscent of the water-filling scheme at high SNR, but the resource allocation is performed at the channel by passive beamforming. *Second,* the impact of passive beamforming on waveform design is subtle. The amplitude of modulated and multisine waveforms with and without IRS are shown in Figs. 7b–7d. Note that no multisine is used for WIT and the case is not shown.

Comment 2.3 — *In the introduction, it is suggested to also discuss the new design challenges posed by the nonlinear model compared to the existing designs under the linear model.*

Response: We believe this is an important issue and have modified the manuscript accordingly as follows.

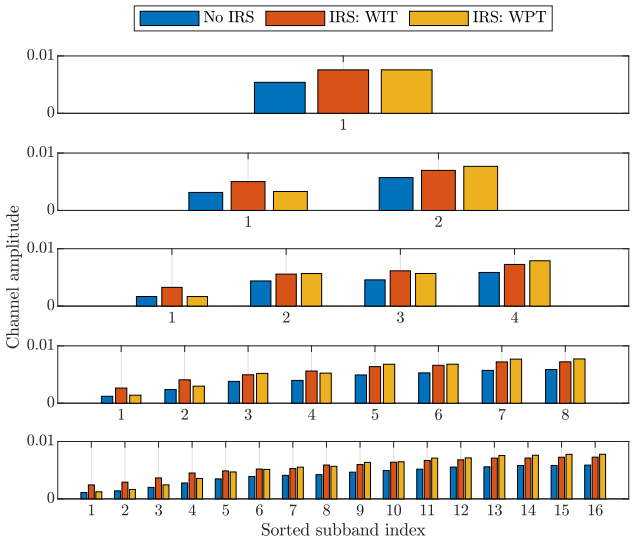
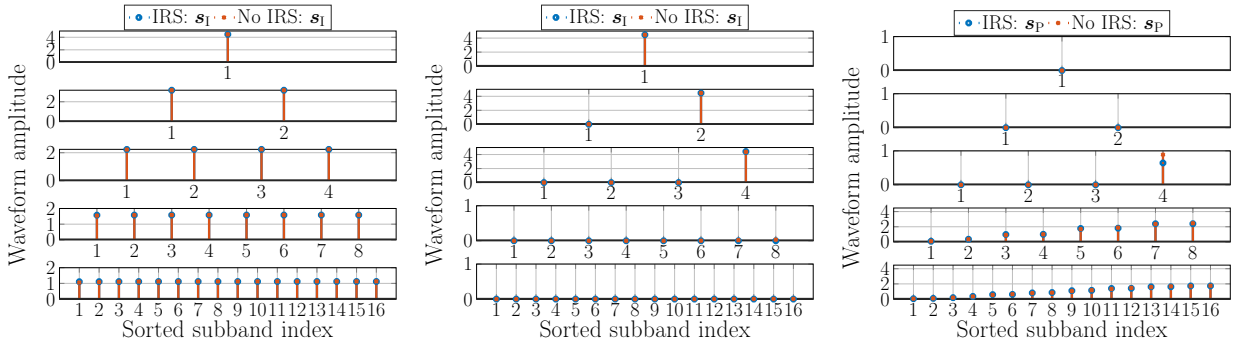


Figure F1: Channel amplitude with and without IRS versus N for $M = 1$, $L = 100$, $\sigma_n^2 = -40\,\text{dBm}$, $B = 10\,\text{MHz}$ and $d_\text{H} = d_\text{V} = 2\,\text{m}$.



(a) WIT waveform: modulated (b) WPT waveform: modulated (c) WPT waveform: multisine

Figure F2: Sorted modulated and multisine amplitude for WIT and WPT with and without IRS versus N for $M = 1$, $L = 20$, $\sigma_n^2 = -40\,\text{dBm}$, $B = 10\,\text{MHz}$ and $d_\text{H} = d_\text{V} = 2\,\text{m}$.

First, we propose a novel IRS-aided SWIPT architecture based on joint waveform, active and passive beamforming design under the diode nonlinear model proposed in [8]. Although this tractable harvester model accurately reveals how the input power level and waveform shape influence the output DC power, it also introduces design challenges such as frequency compensation (i.e., components of different frequencies compensate and produce DC), waveform coupling (i.e., different waveforms jointly contribute to DC), and non-concavity from high-order objective function.

Comment 2.4 — *In 15, the authors integrated the waveform design and active beamforming design into* $\boldsymbol{w}_\text{I}$ *and* $\boldsymbol{w}_\text{P}$*. However, they separately optimized them later. Is it possible to directly optimize* $\boldsymbol{w}_\text{I}$ *and* $\boldsymbol{w}_\text{P}$ *as a whole?*

Response: For the single-user case, direct optimization on $\boldsymbol{w}_{\mathrm{I}}$ and $\boldsymbol{w}_{\mathrm{P}}$ is possible. The reviewer is referred to Algorithm 1 of [14] for more details, and the idea is briefly described as follows. First, $\boldsymbol{w}_{\mathrm{I}}$ and $\boldsymbol{w}_{\mathrm{P}}$ are complex variables of size $MN \times 1$. To enable GP tools (which apply to problems involving real-valued variables), we notice that the optimal phase of $\boldsymbol{w}_{\mathrm{I}}$ and $\boldsymbol{w}_{\mathrm{P}}$ should be designed to compensate the phase of the corresponding composite channel $\boldsymbol{h}$, namely

$$\arg(\boldsymbol{w}_{\mathrm{I}}) = \arg(\boldsymbol{w}_{\mathrm{P}}) = -\arg(\boldsymbol{h}), \tag{E6}$$

such that, from the perspective of phase design, the current expression 9 is maximized (where all cos terms peak at 1) and the rate expression is maximized (where the phases align with those of matched filters). Therefore, the direct optimization on $\boldsymbol{w}_{\mathrm{I}}$ and $\boldsymbol{w}_{\mathrm{P}}$ boils down to amplitude optimization on $\boldsymbol{s}_{\mathrm{I}}$ and $\boldsymbol{s}_{\mathrm{P}}$, where $\boldsymbol{s}_{\mathrm{I/P}} = |\boldsymbol{w}_{\mathrm{I/P}}|$. GP tools can be applied at this stage and the number of real-valued variables are $2MN$. In contrast, the proposed Algorithm 2 decoupled the design in the spatial and frequency domains, derived the global optimal active precoder, and leaves an resource allocation issue with $2N$ real-valued variables in the frequency domain. Since the computational complexity of GP scales exponentially with the number of variables, optimizing $\boldsymbol{w}_{\mathrm{I}}$ and $\boldsymbol{w}_{\mathrm{P}}$ as a whole can be very inefficient and not discussed in the manuscript. Accordingly, we have revised the manuscript as follows.

> To avoid straightforward optimization over complex vectors $\boldsymbol{w}_{\mathrm{I/P}}$ of size $MN \times 1$, we decouple the waveform in the spatial and frequency domains to reduce the size of variables.

Comment 2.5 — *In the low-complexity BCD, it is unclear how the authors dealt with the optimization of ρ and δ. In Algorithm 5, both of them are input. Did the authors perform a two-dimensional search here? Besides, they were assumed to be identical in the simulation. How will this assumption influence the performance loss?*

Response: Thank you for raising this point. The combining ratio δ determines the weight on multisine power waveform at the transmitter, while the splitting ratio ρ determines the priority of the energy harvester at the receiver. We notice that $\delta^\star = \rho^\star = 0$ at the WIT point and $\delta^\star = \rho^\star = 1$ at the WPT point. Heuristically, δ and ρ should approximately equal at each point to boost R-E tradeoff. Indeed, the LC-BCD algorithm intends to draw the R-E tradeoff by performing a two-dimensional search that adjusts δ and ρ from 0 to 1. However, based on the deduction above, we set $\delta = \rho$ is the simulation and obtain the R-E region by one-dimensional search instead. The results show that even a one-dimensional search for the LC-BCD algorithm can achieve a very good performance. We have summarized the point in the manuscript as follows.

> ... where the combining ratio δ determines the weight on multisine power waveform at the transmitter and the splitting ratio ρ determines the priority of the energy harvester at the receiver[2].
>
> ... and the LC-BCD algorithm characterizes the R-E tradeoff via two-dimensional search over δ, ρ from 0 to 1.
>
> ... assume $\delta = \rho$ for simplicity (thus one-dimensional search for the R-E region).
>
> ... *Third,* the LC-BCD algorithm achieves a decent R-E region even if one-dimensional search is considered for $\delta = \rho$ from 0 to 1, which achieves a good balance between performance and complexity.

Comment 2.6 — *If possible, in Figs. 12a and 12b, the authors can show the performance of the design under the conventional linear harvester model, so as to show how much performance loss it would incur.*

[2]We notice that $\delta^\star = \rho^\star = 0$ at the WIT point and $\delta^\star = \rho^\star = 1$ at the WPT point. Heuristically, $\delta^\star$ and $\rho^\star$ should approximately equal at each point to boost R-E tradeoff.

Response: We would like to clarify that the performance loss under the conventional linear harvester model is revealed instead in Figs. 10b and 11b. Figs. 12a and 12b are intended to compare different IRS strategies when both waveform and passive beamforming are designed under nonlinear harvester model. The authors believe that to design passive beamforming under linear harvester model when the waveform is designed under nonlinear harvester model might confuse the readers. On the other hand, we agree with the reviewer that the performance loss under linear harvester deserves more attention, and have changed the label in Figs. 10b and 11b from 'ASS' to 'LEH' and revised the manuscript as follows to make this point clear.

> *Second,* the conventional Linear Energy Harvester (LEH) model leads to power-inefficient design. To show the underlying performance loss, we truncate the DC objective function (8) at $n_0 = 2$ such that, 1) in the passive beamforming problem, $z(\mathbf{\Phi}) = \beta_2 \rho (t_{\mathrm{I},0} + t_{\mathrm{P},0})/2$ and no SCA is required, 2) in the waveform design problem, the WPT-optimal strategy is the Adaptive Single Sinewave (ASS) waveform strategy that allocates all power to the multisine at the strongest subband [8]. As shown in Figs. 10b and 11b, those conventional designs do not exploit the harvester nonlinearity and end up with a nearly 20 dB current gap compared to the nonlinear model-based SMF and GP.

Comment 2.7 — *It was shown via simulation that PS is not always superior to TS, unlike the linear model. Is it possible to develop some analytical preconditions to determine which one is better?*

Response: We concluded that PS outperforms TS when the number of subcarrier is relatively small or when the SNR is relatively large. However, the analytical preconditions in the proposed system can be very hard to obtain because SNR can depend on many factors (e.g., distance of all paths, number of IRS elements, average noise power, etc.). Essentially, this behavior origins from the waveform selection and we would like to analyze it in a nutshell.

TS only performs a naive time-sharing between the WIT point and the WPT point. At the WIT point, only modulated waveform is used to maximize the achievable rate. At the WPT point, the optimal waveform depends on the number of subcarriers N. Although the modulated waveform has an modulation gain of 2 that outstands for a small N, it creates no DC coupling between frequencies and is outperformed by multisine for a large N. Therefore, for a relatively small N, the modulated waveform is used at both WIT and WPT point, and one can heuristically infer that no multisine waveform is needed for any point (this is verified in simulation). In this case, the R-E region is convex, which aligns with the conventional linear harvester model (PS outperforms TS and dedicated power waveform is unnecessary). However, as N becomes sufficiently large, multisine waveform further boosts the WPT point due to its power benefit under nonlinear harvester model. It creates some concavity in the high-power region and accounts for the superiority of TS.

Similar effect exists for SNR. At high SNR, the SWIPT user can achieve a decent rate even if a small fraction of transmit power is allocated to the modulated waveform. Therefore, one can approach high-rate region fast at a small cost of DC power at high SNR. This is however impossible at low SNR.

The manuscript has been updated as follows.

> *Second,* the R-E region is convex for $N \in \{2, 4\}$ and concave-convex for $N \in \{8, 16\}$, such that PS outperforms TS for a small N and is outperformed for a large N. When N is in between, the optimal strategy is a combination of both, i.e., a time sharing between the WPT point and the saddle PS SWIPT point (as denoted by the red curve in Fig. 6a). When N is relatively small, only modulated waveform is used at both WIT and WPT point, and one can deduct that no multisine waveform is needed for the entire R-E region. It aligns with the conventional linear harvester model where the R-E region is convex, PS outperforms TS, and dedicated power waveform is unnecessary. As N becomes sufficiently large, multisine waveform further boosts WPT and creates some concavity in the high-power region, which accounts for the superiority of TS under nonlinear harvester model. Therefore, we conclude that the

rectifier nonlinearity enlarges the R-E region by favoring a different waveform and receiving mode, both heavily depending on N.

Comment 2.8 — *Is it possible to give the complexity order of the algorithms presented in this paper?*

Response: We have added complexity analysis in the revised manuscript. The reviewer is referred to Response 1.1 for details.

Comment 2.9 — *The expressions of signals in (1), (2), and (5) are not conventional. It may be better to give a per-band signal instead of their superposition. Besides, it is better to first introduce the functionalities of the modulated and multisine waveforms before giving (1).*

Response: We agree with the reviewer that the signal model requires more description. The manuscript has been updated as follows.

As suggested in [14], we superpose a multi-carrier modulated information-bearing waveform to a multi-carrier unmodulated power-dedicated waveform (deterministic multisine) to boost the spectrum and energy efficiency. The information signal transmitted over subband $n \in \mathcal{N} \triangleq \{1, \ldots, N\}$ at time t is

$$\boldsymbol{x}_{\mathrm{I},n}(t) = \Re\left\{\boldsymbol{w}_{\mathrm{I},n}\tilde{x}_{\mathrm{I},n}(t)e^{j2\pi f_n t}\right\}, \tag{E7}$$

where $\boldsymbol{w}_{\mathrm{I},n} \in \mathbb{C}^{M\times 1}$ is the information precoder at subband n, $\tilde{x}_{\mathrm{I},n} \sim \mathcal{CN}(0,1)$ is the information symbol at subband n, and f_n is the frequency of subband n. On the other hand, the power signal transmitted over subband n at time t is

$$\boldsymbol{x}_{\mathrm{P},n}(t) = \Re\left\{\boldsymbol{w}_{\mathrm{P},n}e^{j2\pi f_n t}\right\}, \tag{E8}$$

where $\boldsymbol{w}_{\mathrm{P},n} \in \mathbb{C}^{M\times 1}$ is the power precoder at subband n. Therefore, the superposed signal transmitted over all subbands at time t is

$$\boldsymbol{x}(t) = \Re\left\{\sum_{n=1}^{N}(\boldsymbol{w}_{\mathrm{I},n}\tilde{x}_{\mathrm{I},n}(t) + \boldsymbol{w}_{\mathrm{P},n})e^{j2\pi f_n t}\right\}. \tag{E9}$$

Comment 2.10 — *In Remark 2, the meaning of the sentence, "there exists a tradeoff for auxiliary link control in the frequency domain", is not clear to me.*

Response: By 'auxiliary link control in the frequency domain', we intended to emphasize that the auxiliary link cannot be designed independently at different frequencies, and a resource allocation opportunity at the channel arises thanks to the IRS. The composite channel can be tuned flexibly to satisfy the specific requirement of the multi-carrier transmission. For example, one can design the reflection pattern to either enhance the strongest subband (e.g., $\max_{\boldsymbol{\phi},n}\|\boldsymbol{h}_n\|$), or improve fairness among subbands (e.g., $\max_{\boldsymbol{\phi}}\min_n\|\boldsymbol{h}_n\|$). We have added Fig. F1 and updated the Remark 2 to clarify this point.

The cascaded channel varies at different frequencies. Since the reflection cannot be designed independently at different subbands, there exists a tradeoff for the passive beamforming design in the frequency domain. The composite channel can be tuned flexibly to meet the specific requirement of multi-carrier SWIPT. For example, one can design the reflection pattern to either enhance the strongest subband

(e.g., $\max_{\boldsymbol{\phi},n}\|\boldsymbol{h}_n\|$), or improve fairness among subbands (e.g., $\max_{\boldsymbol{\phi}}\min_n\|\boldsymbol{h}_n\|$). This is essentially a resource allocation opportunity at the channel enabled by the IRS. In the MISO case, a similar effect also exists in the spatial domain. Therefore, each reflection coefficient is indeed shared by M antennas over N subbands.

Comment 2.11 — *Some editorial comments: 1) "a (->an) M-antenna"; 2) "through a (->an) L-element IRS"; 3) "the CSIT of direct and cascaded channels are (->is) known"; 4) "It demonstrated (->demonstrates) that"; 5) In Figs. 10b and 11b, the text in the y-axis is too long. Please divide it into two rows. Besides, the text in the x-axis of the upper figure is covered by the lower one.*

Response: We sincerely appreciate your careful reading and rigorous feedbacks. The manuscript have been revised accordingly and we will endeavor to avoid similar issues in the future.

Reviewer 3

Comment 3.1 — *Proposition 5 is questionable. The algorithm may not converge to a local optimal due to the coupling constraint* (15b).

Response: Thank you for pointing out this issue. Indeed, the local optimality is not guaranteed as (15b) couples all blocks together. We have removed the claim on optimality and updated Proposition 5 as follows.

> For any feasible initial point, the BCD Algorithm 4 is guaranteed to converge.

The proof has been modified correspondingly.

> The objective function (15a) is non-decreasing over the iterations, because the IRS phase shift by Algorithm 1 is local optimal, the active precoder by (27) is global optimal, and the waveform amplitude and splitting ratio by Algorithm 2 are local optimal. (15a) is also upper-bounded due to the unit-modulus constraint (15d) and the transmit power constraint (15c). Therefore, Algorithm 4 is guaranteed to converge. However, it may not converge to optimal points since variables are coupled in constraint (15b) [58].

Comment 3.2 — *The figures are too small. For example, Fig. 11b, the axis legend is not fully shown.*

Response: We apologize for the inconvenience. The figures have been amplified and the legends have been calibrated.

Comment 3.3 — *It is said that the reference path loss is* 35 dB *at* 1 m*, while for the center frequency at* 5.18 GHz*, even adopting the free space path loss model also will lead a much sever path loss at* 1 m*. It is not clear how the authors calculate this.*

Response: We agree with the reviewer that the reference path loss used in the simulation is faulty and we sincerely apologize for this issue. The parameter 35 dB was considered in some IRS literatures without specifying the operating frequency, and it was taken in the previous simulations without careful consideration.

In the corrected manuscript, we instead consider a center frequency at 2.4 GHz that corresponds to a free-space path loss of 40 dB at 1 m. The path loss used in the simulation is taken from IEEE TGn channel model D [48] where the slope is set to 2 (i.e., free-space model) up to a breakpoint distance of 10 m, and set to 3.5 onwards. This has a similar effect to most IRS literatures that assign a relatively large path loss exponent to the direct path and smaller ones to the incident and reflected paths to penalize the channel with large distance.

In the updated manuscript, we raise the EIRP from 36 dBm to 40 dBm (which still complies with FCC regulations [R2]) and change the receive antenna gain from 2 dBi to 3 dBi to compensate the 5 dB difference. Please refer to the following specification for details.

> We consider a large open space Wi-Fi-like environment at 2.4 GHz center frequency where the channel is modeled by IEEE TGn channel model D [48]. Specifically, the path loss exponent is set to 2 (i.e., free-space model) up to a breakpoint distance of 10 m, and set to 3.5 onwards. All fadings are modeled as Non-Line-of-Sight (NLoS) with tap delays and powers specified in model D. The tap gains are modeled as i.i.d. CSCG variables. Rectenna parameters are set to $k_2 = 0.0034$, $k_4 = 0.3829$, $R_{\mathrm{A}} = 50\,\Omega$ [8] such that $\beta_2 = 0.17$ and $\beta_4 = 957.25$. We also choose the average Effective Isotropic Radiated Power (EIRP) as $P = 40$ dBm, the receive antenna gain as 3 dBi, the scaling ratio as $\alpha = 2$, the tolerance as $\epsilon = 10^{-8}$, and assume $\delta = \rho$ for simplicity (thus one-dimensional search for the R-E region).

Comment 3.4 — *The authors considered the practical reflection coefficient beamforming in the simulation results. It is not clear how the authors obtain the discrete phase shift results. Direct quantization method and some other customized optimization techniques for discrete phase shifts may have significant performance gap, see [R3]. Also, this work adopts the same uniform quantization as the above work, which thus needs to be clarified in the paper.*

Response: Thank you for the clarification. The discrete phase shift is obtained by quantizing the continuous phase shift obtained by Algorithm 4 to the uniform IRS codebook $\mathcal{C}_\phi = \{e^{j2\pi i/2^b} \mid i = 1, \ldots, 2^b\}$. We are aware that quantization can lead to performance loss compared with direct optimization over the feasible phase shift set. In the revised manuscript, we have emphasized this point as follows.

> On the other hand, since the practical reflection coefficient depends on the available element impedances, we consider a discrete uniform IRS codebook $\mathcal{C}_\phi = \{e^{j2\pi i/2^b} \mid i = 1, \ldots, 2^b\}$ and perform quantization on the continuous reflection coefficients returned by Algorithm 4 to reduce the circuit complexity and control overhead[3].

Comment 3.5 — *There are also some other practical phase shift models and nonlinear energy harvesting models such as saturation model for IRS-aided SWIPT in existing works, which may be discussed as related works such as amplitude-dependent phase shift model.*

Response: We have added reference [R1] in the updated manuscript that models the saturation effect of energy harvesters and physical properties of the IRS. Its contributions have been summarized in the manuscript as follows.

[3]Note that this relax-then-quantize approach can bring notable performance loss compared with direct optimization over the discrete phase shift set, especially for a small b (i.e., low-resolution IRS). The readers are referred to [R3] for details.

In [R1], the authors proposed a scalable resource allocation framework for SWIPT systems involving large-scale IRS, where the reflectors are grouped into tiles and the optimization process is divided into an offline mode-design stage and an online mode-selection stage to reduce overall complexity. It was concluded that the tile-based two-stage algorithm not only enables a flexible balance between performance and complexity but also adapts well to practical IRS and harvester models.

The future works has been updated as follows.

Also, harvester saturation effect and more involved IRS model based on direction- and frequency-dependent reflection [24] and/or partially/fully-connected reflection [53] could be considered in future works.

Comment 3.6 — *Some early magazine papers and recent tutorial papers related to IRS-aided WPT/WIT may be further discussed.*

Response: We have incorporated those references into the manuscript as follows.

... The hardware architecture, design challenges and application opportunities of practical IRS are covered in [R4].

... A recent tutorial paper [R5] provided a comprehensive overview on IRS-aided wireless-powered networks, where channel estimation, resource allocation and practical constraints are discussed in detail.

Reviewer 4

Comment 4.1 — *What is the computational complexity of the proposed algorithm?*

Response: We have added complexity analysis in the revised manuscript. The reviewer is referred to Response 1.1 for details.

Comment 4.2 — *Some related works are worth citing, such as [R6].*

Response: We have incorporated the reference into the manuscript as follows.

Deep reinforcement learning tools were also applied in [R6] to assist practical secure beamforming and reflection pattern design under QoS constraints for time-varying channels.

References

[R1] D. Xu, V. Jamali, X. Yu, D. W. K. Ng, and R. Schober, "Optimal resource allocation design for large irs-assisted swipt systems: A scalable optimization framework," *arXiv preprint arXiv:2104.03346*, pp. 1–30, 2021.

[R2] AFAR Communications, "Fcc rules for unlicensed wireless equipment operating in the ism bands." [Online]. Available: https://afar.net/tutorials/fcc-rules/

[R3] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1838–1851, Mar. 2020.

[R4] ——, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, Jan. 2020.

[R5] Q. Wu, X. Guan, and R. Zhang, "Intelligent reflecting surface aided wireless energy and information transmission: An overview," *arXiv preprint arXiv:2106.07997*, pp. 1–39, Jun. 2021.

[R6] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 375–388, 2021.