
Fudong Li, Hai Jiang, Senior Member, IEEE, Rongfei Fan, Member, IEEE, and Peng Tan, Member, IEEE

Abstract—We consider a cognitive radio system, in which a secondary transmitter harvests energy from a primary transmitter’s RF signal. The secondary transmitter, which provides decode-and-forward relaying service for the primary system, transmits its own data by using downlink non-orthogonal multiple access (NOMA). A time-switching protocol is used by the secondary transmitter to harvest energy and decode the primary transmitter’s information. Our objective is to achieve maximal secondary throughput, by optimally selecting the time portion used for energy harvesting and the secondary transmitter’s power allocation in NOMA transmission. Two optimization problems are formulated, in which the secondary receiver performs or does not perform successive interference cancellation (SIC), respectively. Although the two problems are nonconvex, we devise a method to transform the problems into equivalent problems under difference cases. Then we theoretically prove that the objective functions of the equivalent problems are quasiconcave, based on which we develop two-level bisection search algorithms to solve the equivalent problems. Interestingly, we show that the secondary receiver does not always guarantee a higher secondary throughput than the case without performing SIC. Computer simulation demonstrates that our algorithm performs better than an equal power allocation algorithm and an orthogonal multiple access algorithm.

Index Terms—Cognitive radio, cooperation, relay, decode-and-forward, non-orthogonal multiple access, energy harvesting.

I. INTRODUCTION

Communications over wireless channels are technically restricted by the bottleneck of spectrum efficiency, on which research efforts have been made for decades, and many state-of-the-art techniques have been devised. Among the techniques, cognitive radio (CR) and non-orthogonal multiple access (NOMA) are quite promising. CR alleviates the conflict between explosively increased mobile applications and scarce frequency resources, and significantly improves the spectrum efficiency mainly in three ways of spectrum access: interweave, underlay, and overlay [2]–[6]. As a novel paradigm of transmission, NOMA is recognized to be a key enabling technology to further improve spectrum efficiency in 5G and beyond [7]–[12]. NOMA does not have the limitation of transmitting only one user’s signal in one resource block over traditional orthogonal multiple access (OMA) systems. The power domain, treated as a new degree of freedom of resource management, is inherently exploited in NOMA systems. Specifically, at the transmitter side, different users’ signals are assigned with distinctive power levels to enable concurrent transmissions. At the receiver side, successive interference cancellation (SIC) is adopted to perform user information recovery.

Naturally, the interaction of NOMA with CR, termed cognitive NOMA, yields even more spectrum-efficient communications [13]–[20]. The CR inspired NOMA is conceptually introduced in [13], in which two users with different priorities are paired on one resource block. The weak user who undergoes poor channel gain from the transmitter has a higher priority, and thus, is deemed as the primary user (PU) with a target reception quality. The strong user who experiences good channel gain from the transmitter has a lower priority, and thus, is deemed as the secondary user (SU). Under the condition that the PU’s target reception quality is guaranteed, the transmitter uses NOMA to send simultaneously to the PU and SU. The work in [14] extends the CR inspired NOMA concept to the scenario with multiple PUs and one single SU, in which the SU, having better channel gain than those of the PUs, is served together with all the PUs simultaneously by using NOMA. The transmission strategy of the multiple-antenna transmitter is designed such that the energy efficiency is maximized. By further extending CR inspired NOMA with multiple PUs and multiple SUs, the work in [15] proposes a distributed matching method, which pairs an SU with a PU and assigns transmit power levels for them, targeting system throughput maximization. In [16], underlaid with the primary network, the secondary network sets up an interference guard zone to limit the primary interference, and transmits information from one secondary transmitter (ST) to multiple secondary receivers (SRs) by employing NOMA. For secondary transmissions, the outage performance and diversity order are analyzed. A two-user underlay cognitive NOMA system is investigated in [17], in which a dedicated full-duplex relay is resorted to help a far user to forward information from the base station (BS). The power allocation, beamformer design and the outage performance are investigated. The work in [18] considers an overlay cognitive NOMA system, in which
the primary network is assisted by the secondary network that applies NOMA principle in its transmissions. In the first phase, a primary transmitter (PT) transmits signals to its primary receiver (PR) and an ST, while in the second phase, the ST cooperatively transmits the PR’s and the SRs’ signals utilizing NOMA. The outage performance of the primary and secondary systems are characterized. The work in [19] considers an overlay cognitive NOMA system with a number of STs, in which one ST is scheduled to use NOMA to forward primary signals and send its own signals to its receivers. Two scheduling schemes, targeting reliability and fairness, respectively, are proposed. A survey of cognitive NOMA techniques and future research trends is given in [20].

Energy efficiency is another major concern in wireless communications. Conventional wireless devices, powered by batteries, may experience energy depletion and inconvenience of battery replacement. Energy harvesting provides a promising and viable solution [21], [22], in which energy from various sources is harvested and used for transmissions. Apart from natural energy sources (such as wind and solar), wireless RF signal is a widely studied energy source since it can be obtained almost everywhere and at any time. Moreover, the simultaneous wireless information and power transfer (SWIPT) technique has the advantage of transferring energy and delivering information simultaneously [22]–[24]. To achieve SWIPT, two typical implementation protocols, time-switching (TS) and power-splitting (PS), are designed [24].

By integrating SWIPT into cognitive NOMA framework, a greener and more sustainable communication is expected. In the literature, some research efforts on SWIPT-based cognitive NOMA are conducted, such as [25], [26]. The work in [25] studies an underlay cognitive NOMA system, in which multiple STs harvest energy from one energy transmitter and transmit to one common BS by using NOMA. The secure energy efficiency maximization problem for secondary system is investigated. The work in [26] can be viewed as a CR inspired NOMA system enhanced with SWIPT. To serve a weak user, a strong user uses PS protocol to harvest energy from the transmitter’s signals and decode its own and the weak user’s information during the first phase. Afterwards, during the second phase, the strong user forwards the weak user’s information. With a condition that the reception quality of the weak user is satisfied, the strong user’s throughput is maximized, by optimally designing the beamformers of the multiple-antenna transmitter and the PS ratio at the strong user.

In this paper, we investigate an overlay cognitive NOMA scenario equipped with SWIPT, in which an ST uses TS protocol for energy harvesting and information decoding, and uses NOMA to help a PT and deliver its own information. The optimal solution, deriving the time ratio for energy harvesting as well as the NOMA power allocation of the ST, is provided in our paper. Our main contributions are summarized as follows. 1) A new SWIPT-enhanced cognitive NOMA framework: In the literature, SWIPT is integrated with underlay cognitive NOMA in [25], and with CR-inspired NOMA in [26]. However, SWIPT-enhanced overlay cognitive NOMA is not investigated in the literature, and its performance is unclear. To address this research gap, we investigate a SWIPT-enhanced overlay cognitive NOMA framework. 2) Optimal solution: The formulated problems are nonconvex, and are generally hard to solve. We devise a method to transform the formulated problems to equivalent problems under different cases. We theoretically prove that the objective functions of the equivalent problems are quasiconcave. We then develop an effective algorithm, by using a two-level bisection search, to find the optimal solution of each equivalent problem. We also develop a method that could reduce the number of iterations in the inner bisection search. 3) Interesting insights: Interestingly, different from existing NOMA works in which SIC is always applied, performing SIC in our work does not guarantee a better performance than the scenario without performing SIC. The insight behind this observation is also discussed.

II. SYSTEM MODEL

We consider a cognitive system, consisting of one pair of PT and PR, and one pair of ST and SR, as depicted in Fig. 1. The spectrum is licensed to the primary system. The PT has a stable power supply, which transmits data with a fixed transmit power $P_p$. The ST is powered by harvested energy from RF signals transmitted by the PT. The ST opportunistically gains spectrum access opportunities in an overlay mode, i.e., when the link from the PT to the PR is not good enough, the ST could help forward the PT’s signal and send its own signal to the SR as well by using NOMA.

The system is time slotted, and each time slot has a unit length. The channel coefficients between PT and PR, PT and SR, ST and PR, and ST and SR are denoted as $h_{pp}$, $h_{ps}$, $h_{sp}$, and $h_{ss}$, respectively. In subscript of the channel coefficients, the first symbol $p$ or $s$ means the primary or secondary transmitter, and the second symbol $p$ or $s$ means the primary or secondary receiver. In addition, the channel coefficient between PT and ST is denoted as $h_{ts}$. Accordingly, the channel gains (square of channel coefficient magnitude) of those channels are denoted as $g_{pp}$, $g_{ps}$, $g_{sp}$, $g_{ss}$, and $g_{tt}$, respectively. Block fading is assumed, which means that all the channel gains keep unchanged in each time slot and may change independently from slot to slot.
Basically, in a particular time slot, the system works in one of two transmission modes:

- Direct primary transmission mode
- Cooperative transmission mode.

### A. Direct Primary Transmission Mode

We assume that the PT and ST always have enough information to be transmitted to the PR and SR, respectively. The PT has a target throughput, denoted as $\gamma^T$, for each time slot. At a slot, the achievable throughput of the direct link from the PT to the PR is expressed as $\log(1 + \frac{P_g h}{\sigma^2})$, where $\sigma^2$ is variance of the background additive white Gaussian noise (AWGN).

If the achievable throughput is not less than the PT’s target throughput $\gamma^T$ (which equivalently means that the channel gain of the direct primary link $P_T \rightarrow P_R$ satisfies $g_{pp} \geq (2\gamma^T - 1)\sigma^2/P_p$), then the system works in the **direct primary transmission mode**. In this mode, the PT transmits data to the PR during the whole slot, whereas the ST is not allowed to access the spectrum. Thus, the ST harvests energy during the whole slot, which yields the battery energy at the end of the slot as $\max\{E_0 + \eta P_g h, 0\}$, where $E_0$ is the ST’s battery energy at the beginning of the slot, $\eta \in (0, 1)$ is the energy conversion efficiency, and $E_c$ is the energy expenditure in a time slot due to circuit operation and channel estimation.

### B. Cooperative Transmission Mode

If the achievable throughput $\log(1 + \frac{P_g h}{\sigma^2})$ of the direct primary link $P_T \rightarrow P_R$ is lower than the PT’s target throughput $\gamma^T$ (which equivalently means that $g_{pp} < (2\gamma^T - 1)\sigma^2/P_p$), the ST is requested to help in forwarding the PT’s message to PR by using decode-and-forward (DF) relaying. The ST can simultaneously send its own message to the SR based on downlink NOMA. In other words, the system works in the **cooperative transmission mode**.

In this mode, the slot is partitioned into three phases, as shown in Fig. 2. The first phase has length $\rho \in [0, 1]$, while the second and third phases both have length $(1 - \rho)/2$, where the value of $\rho$ is a designed parameter, to be optimized hereafter.

1) **The First Phase**: During the first phase, the PT transmits wireless RF signals with power $P_p$, from which the ST harvests energy. Then the ST’s battery energy, after harvesting, has the form of $E_h = E_0 + \eta P_g h$. If $E_h$ is less than $E_c$, then the ST is not able to help the PT, and thus, the system has to work in the direct primary transmission mode. Therefore, for the system to work in the cooperative transmission mode, we should have $E_h \geq E_c$, which leads to $\rho \geq \max\{\frac{E_c - E_0}{\eta P_g h}, 0\}$.

2) **The Second Phase**: During the second phase, the PT transmits its information signal $x_p$ with power $P_p$, which is received by the PR and ST, as well as the SR.

At the PR, the received signal in the second phase of the slot is represented by $\sqrt{P_p h_{pp} x_p + n_{pr}}$, in which $n_{pr}$ is the AWGN at the PR.

The received signal at the ST in the second phase of the slot is written as $\sqrt{P_p h_{tt} x_p + n_{st}}$, in which $n_{st}$ is the AWGN at the ST. The achievable information rate for the transmission from the PT to the ST in the second phase of the slot is expressed as $R_{tt} = \frac{1 - \rho}{2} \log(1 + \frac{P_g h}{\sigma^2})$. As the ST needs to decode the PT’s signal, the achievable information rate $R_{tt}$ should not be less than the target throughput $\gamma^T$ of the primary system. Thus, we should have constraint $R_{tt} \geq \gamma^T$, based on which we have constraint $\rho \in \mathcal{A}_1 \triangleq \left[ \max\{\frac{E_c - E_0}{\eta P_g h}, 0\}, 1 - \frac{2\gamma^T}{\log(1 + \frac{P_g h}{\sigma^2})} \right]$.

At the SR, the received signal in the second phase of the slot is represented by $\sqrt{P_p h_{ps} x_p + n_{sr}}$, in which $n_{sr}$ is the AWGN at the SR.

3) **The Third Phase**: During the third phase, the PT transmits a copy of the primary signal $x_p$ with power $P_p$. The ST applies downlink NOMA to transmit a superimposed signal consisting of the PT’s signal $x_p$ and the ST’s own signal $x_s$, by using the harvested energy in a greedy manner, i.e., it uses up all the available energy stored in battery in this phase. In specific, the ST transmits $x_{st} = \sqrt{\alpha P_p x_p} + \sqrt{(1 - \alpha)} P_s x_s$, where $\alpha$, a parameter to be optimized, is the power ratio for $x_p, (1 - \alpha)$ is the power ratio for $x_s$, and $P_s$ is the transmit power of ST, shown as $P_s = \frac{2(E_h - E_c)}{1 - \rho} = \frac{2(\eta P_g h + E_0 - E_c)}{1 - \rho}$. (1)

Accordingly, the received signal at the PR in the third phase is $\sqrt{P_p h_{pp} x_p + h_{sp} x_{st} + n_{pr}}$. As the PR receives the PT’s signal $x_p$ in the second phase (from the PT) and the third phase (from the PT and ST), the PR employs maximal ratio combining (MRC) to combine the received PT’s signal portions, which yields the overall throughput of the PT’s signal $x_p$ (also called the throughput of the primary system) in the time slot as $R_{pp} = \frac{1 - \rho}{2} \log \left(1 + \frac{P_g h_{pp}}{\sigma^2} + \frac{P_g h_{pp} + \alpha P_s}{(1 - \alpha)(P_c g_{sp} + \sigma^2)}\right)$. (2)

1Note that in the first phase, the wireless RF signals can actually be information signals of the PT for the PR. In other words, in the first phase, from the PT’s signals, the ST tries to harvest energy while the PR tries to decode information. This portion of information for the PR will be investigated in Section III-F.

2Note that our method can be straightforwardly extended to the case when there exists a limit for the energy level that the ST can use in the third phase of the slot (i.e., there is some energy left at the end of the slot).
In the third phase, the received signal at the SR is \( y_{ss} = \sqrt{P_p h_p x_p + h_{ss} x_{st} + n_{sr}} \). In the received signal \( y_{ss} \), the PT’s signal portion \( x_p \) is interference to the SR’s own signal \( x_s \). Thus, similar to NOMA works in the literature, the SR can use SIC, i.e., it first decodes the PT’s signal \( x_p \), removes \( x_p \) from \( y_{ss} \), and then decodes its own signal \( x_s \).

When the SR decodes the PT’s signal portion \( x_p \), the achievable information rate is given as

\[
R_{ps} = \frac{1 - \rho}{2} \log \left( 1 + \frac{P_p g_{ps} + \alpha P_c g_{ss}}{(1 - \alpha) P_c g_{ss} + \sigma^2} \right),
\]

where MRC is used to combine the received PT’s signal portions in the second and third phases. When performing SIC at the SR, \( \rho \) and \( \alpha \) should be set up such that \( R_{ps} \geq \gamma^T \), which means that the SR can successfully decode \( x_p \), and perform SIC to remove \( x_p \) in \( y_{ss} \). Subsequently, the achievable throughput of the secondary system’s throughput (also called the secondary system’s throughput) can be written as

\[
R_{ss} = \frac{1 - \rho}{2} \log \left( 1 + \frac{(1 - \alpha) P_c g_{ss}}{\sigma^2} \right). \tag{4}
\]

Our objective is to maximize the secondary system’s throughput while guaranteeing that the throughput of the primary system is not smaller than \( \gamma^T \). Thus we formulate the following optimization problem.

**Problem P1:**

\[
\begin{align*}
\max_{\alpha, \rho} & \quad R_{ss} = \frac{1 - \rho}{2} \log \left( 1 + \frac{(1 - \alpha) P_c g_{ss}}{\sigma^2} \right) \tag{5a} \\
\text{s.t.} & \quad \rho \in A_1; \min\{R_{pp}, R_{ps}\} \geq \gamma^T; 0 \leq \alpha \leq 1. \tag{5b}
\end{align*}
\]

In Problem P1, to perform SIC at the SR, we have constraint \( R_{ps} \geq \gamma^T \). If the SR does not perform SIC, interference from the PT’s signal is not cancelled, which harms the secondary throughput. On the other hand, if the SR does not perform SIC, the constraint \( R_{ps} \geq \gamma^T \) can be removed, and we can have a larger feasible region of \( \alpha \) and \( \rho \), which benefits the secondary throughput. It is not clear whether the overall effect of not performing SIC is beneficial or harmful. Thus, we should also investigate an optimization problem, in which the SR does not perform SIC.

When the SR does not perform SIC, the achievable throughput of the secondary system is

\[
R_{ss}^{\prime/\prime} = \frac{1 - \rho}{2} \log \left( 1 + \frac{(1 - \alpha) P_c g_{ss}}{(1 - \alpha) P_c g_{ss} + P_p g_{ps} + \sigma^2} \right), \tag{6}
\]

in which superscript \( (\prime/\prime) \) stands for “without performing SIC.” Accordingly, the following optimization problem can be formulated.

**Problem P2:**

\[
\begin{align*}
\max_{\alpha, \rho} & \quad R_{ss}^{\prime/\prime} = \frac{1 - \rho}{2} \log \left( 1 + \frac{(1 - \alpha) P_c g_{ss}}{(1 - \alpha) P_c g_{ss} + P_p g_{ps} + \sigma^2} \right) \tag{7a} \\
\text{s.t.} & \quad \rho \in A_1; R_{pp} \geq \gamma^T; 0 \leq \alpha \leq 1. \tag{7b}
\end{align*}
\]

As a summary, the optimal solution of the system is the better one between the optimal solutions of Problems P1 and P2, which has larger secondary throughput.\(^5\)

III. OPTIMAL SOLUTION OF PROBLEM P1

Problem P1 is non-convex since the objective and the constraint functions are not jointly concave. Generally, it is hard to solve such a problem. In this work, we provide an efficient method to solve our Problem P1.

By careful inspection, there exists the following useful lemma.

**Lemma 1:** When the optimality of Problem P1 is achieved, constraint \( \min\{R_{pp}, R_{ps}\} \geq \gamma^T \) should be active, i.e., we should have \( \min\{R_{pp}, R_{ps}\} = \gamma^T \).

**Proof:** We use proof by contradiction. Suppose that when the optimality of Problem P1 is achieved, we have \( \min\{R_{pp}, R_{ps}\} > \gamma^T \).

From (2) and (3), we know that \( R_{pp} \) and \( R_{ps} \) both are increasing functions of \( \alpha \). Furthermore, when \( \alpha = 0 \), we have \( \min\{R_{pp}, R_{ps}\} = \min\{R_{pp} g_{pp} + P_p g_{pp}, P_p g_{pp} + \sigma^2\} < \gamma^T \), in which the last inequality comes from the fact that

\[
\frac{1 - \rho}{2} \log \left( 1 + \frac{P_p g_{pp} + \sigma^2}{\sigma^2} \right) \leq \frac{1 - \rho}{2} \log \left( 1 + \frac{P_p g_{pp}}{\sigma^2} \right) < \gamma^T.
\]

In (8), the last inequality comes from the fact that the system works in the cooperative transmission mode when the achievable throughput \( \log(1 + P_p g_{pp}/\sigma^2) \) of the direct primary link from the PT to the PR is less than the PT’s target throughput \( \gamma^T \).

Thus, from the optimality point of Problem P1, we can decrease the value of \( \alpha \) such that we still have \( \min\{R_{pp}, R_{ps}\} \geq \gamma^T \) but we have a larger objective function \( R_{ss} \) (noting that \( R_{ss} \) is a decreasing function of \( \alpha \) according to (4)). This is a contradiction.

This completes the proof. \qed

Lemma 1 indicates that, at optimality of Problem P1, either \( R_{pp} \) or \( R_{ps} \) should be equal to \( \gamma^T \).

If \( R_{pp} = \gamma^T \), then from (2), we can see that \( \alpha \) can be expressed by a function of \( \rho \) as

\[
\alpha = \frac{P_p g_{pp} + \sigma^2}{(\mu_p + 1) P_p g_{sp} g_{pp} + \sigma^2} \tag{9}
\]

with

\[
\mu_p \triangleq \frac{P_p g_{pp}}{\sigma^2} - 1. \tag{10}
\]

Note that by using (8), we can see \( \mu_p > 0 \).

When \( \alpha \) is expressed as a function of \( \rho \) as in (9), the constraint \( 0 \leq \alpha \leq 1 \) should be satisfied. Note that \( \alpha > 0 \) is satisfied automatically, since from (8), we have \( \frac{1 - \rho}{2} \log(1 + \frac{2 P_p g_{pp}}{\sigma^2}) < \gamma^T \), based on which we have \( \mu_p \sigma^2 - P_p g_{pp} > 0 \).

\(^5\) If both Problems P1 and P2 are infeasible, which means the cooperative transmission mode cannot make the throughput of the primary system be at least \( \gamma^T \), then the system will work in direct primary transmission mode.
yielding \( \alpha > 0 \). Thus, only \( \alpha \leq 1 \) is considered here, which is equivalent to

\[
\rho \in A_2 \triangleq \{ \rho | \rho \geq 0 \text{ and } P_e g_{sp} - \mu_p \sigma^2 + P_g g_{pp} \geq 0 \}. \quad (11)
\]

**Lemma 2:** Set \( A_2 \) is a closed interval of \( \rho \), with closed-form starting and ending points.

**Proof:** The proof is similar to the proof in Appendix A in [1] (conference version of this paper), and thus, is omitted due to space limit.

If \( R_{pp} = \gamma^T \), then from (3), we can see that \( \alpha \) can be expressed as a function of \( \rho \), as

\[
\alpha = F_e(\rho) \triangleq \frac{\mu_s(P_e g_{ss} + \sigma^2) - P_g g_{ps}}{(\mu_s + 1) P_e g_{ss}} \quad (12)
\]

with \( \mu_s \triangleq \frac{2 \gamma^T}{\sigma^2} - \frac{P_g g_{ps}}{\sigma^2} - 1 \).

When \( \alpha \) is expressed as a function of \( \rho \) as in (12), the constraint \( 0 \leq \alpha \leq 1 \) is equivalent to

\[
\rho \in A_3 \triangleq \{ \rho | \rho \geq 0 \text{ and } \mu_s(P_e g_{ss} + \sigma^2) - P_g g_{ps} \geq 0 \} \cap \{ \rho | \rho \geq 0 \text{ and } P_e g_{ss} - \mu_s \sigma^2 + P_g g_{ps} \geq 0 \}. \quad (13)
\]

In \( A_3 \), set \( \{ \rho | \rho \geq 0 \text{ and } \mu_s(P_e g_{ss} + \sigma^2) - P_g g_{ps} \geq 0 \} \) comes from \( \alpha \geq 0 \), which is a closed interval, since \( \mu_s(P_e g_{ss} + \sigma^2) - P_g g_{ps} \) is an increasing function of \( \rho \). Similar to set \( A_2 \), set \( \{ \rho | \rho \geq 0 \text{ and } P_e g_{ss} - \mu_s \sigma^2 + P_g g_{ps} \geq 0 \} \) is also a closed interval. Thus, \( A_3 \) is a closed interval.

Afterwards, by comparing (2) and (3), we have the following four cases for Problem P1.

A. **Case 1:** When \( g_{pp} \leq g_{ps} \) and \( g_{sp} \leq g_{ss} \)

For a specific value of \( \rho \), \( R_{pp} \) in (2) and \( R_{ps} \) in (3) can be viewed as functions of \( \alpha \).

1) **Intersections of Curves** \( R_{pp} \) vs. \( \alpha \) and \( R_{ps} \) vs. \( \alpha \) over \( \alpha \in [0, 1] \): By considering two curves: \( R_{pp} \) vs. \( \alpha \) and \( R_{ps} \) vs. \( \alpha \) over \( \alpha \in [0, 1] \), we have the following lemma.

**Lemma 3:** For a specific value of \( \rho \), the two curves \( R_{pp} \) vs. \( \alpha \) and \( R_{ps} \) vs. \( \alpha \) over \( \alpha \in [0, 1] \) have up to two intersections.

**Proof:** For a given \( \rho \), from (2) and (3), if we set \( R_{pp} = R_{ps} \), we can obtain

\[
L_1 P_e^2 (1 - \alpha)^2 + \left[ L_1 \left( \frac{\sigma^2_{g_{ss}}}{g_{ps}} + \frac{\sigma^2_{g_{ps}}}{g_{ss}} \right) + \frac{P_g g_{ps} - \mu_s \sigma^2}{g_{ss} g_{sp}} \frac{g_{ps}}{g_{ss}} - \frac{P_g g_{ps}}{g_{ss} g_{sp}} \right]
+ 1 \right] P_e (1 - \alpha) + 2 L_1 \frac{\sigma^4}{g_{ss} g_{sp} g_{ps}} - P_e = 0, \quad (14)
\]

where \( L_1 = \frac{P_g g_{ps} - \mu_s \sigma^2}{g_{ss} g_{sp}} < 0 \). Considering the left hand-side of (14) as a function of \( \alpha \), the two roots of (14) are

\[
\text{ROOT}_1 = 1 - \frac{-L_2 + \sqrt{L_2^2 - 4L_1 L_3}}{2L_1 P_e}, \quad (15)
\]

\[
\text{ROOT}_2 = 1 - \frac{-L_2 - \sqrt{L_2^2 - 4L_1 L_3}}{2L_1 P_e}, \quad (16)
\]

in which \( L_2 = L_1 \left( \frac{\sigma^2_{g_{ss}}}{g_{ps}} + \frac{\sigma^2_{g_{ps}}}{g_{ss}} \right) + \frac{P_g g_{ps} - \mu_s \sigma^2}{g_{ss} g_{sp}} \frac{g_{ps}}{g_{ss}} + 1 \) and \( L_3 = 2L_1 \frac{\sigma^4}{g_{ss} g_{sp} g_{ps}} - P_e < 0 \). Note that when \( L_2 < 0 \), both ROOT$_1$ and ROOT$_2$ are larger than 1 and infeasible.

To make ROOT$_1$ shown in (15) feasible, we should have \( 0 \leq \text{ROOT}_1 \leq 1 \), which equivalently means that \( \rho \in A_4 \triangleq \{ \rho | 0 \leq \rho \leq 1, L_2 \geq 0, L_2 - 4L_1 L_3 \geq 0, 2L_1 P_e + L_2 - \sqrt{L_2^2 - 4L_1 L_3} \leq 0 \} \). The set \( A_4 \) is a closed interval of \( \rho \).

To make ROOT$_2$ shown in (16) feasible, we should have \( 0 \leq \text{ROOT}_2 \leq 1 \), which equivalently means \( \rho \in A_5 \triangleq \{ \rho | 0 \leq \rho \leq 1, L_2 \geq 0, L_2 - 4L_1 L_3 \geq 0, 2L_1 P_e + L_2 + \sqrt{L_2^2 - 4L_1 L_3} \leq 0 \} \). The set \( A_5 \) is an interval of \( \rho \), and is a subset of \( A_4 \).

Overall, for any specific \( \rho \), if \( \rho \in \bar{A}_1 \triangleq [0, 1] \) \( \setminus \mathcal{A}_4 \), then the two curves \( R_{pp} \) vs. \( \alpha \) and \( R_{ps} \) vs. \( \alpha \) do not have intersection; if \( \rho \in A_1 \setminus A_5 \), then the two curves have one intersection denoted as \( (\alpha_{eq, 1}, \text{ROOT}_1) \), with \( \alpha_{eq,1} = \text{ROOT}_1 \); if \( \rho \in A_5 \), then the two curves have two intersections denoted as \( (\alpha_{eq,1}, \text{ROOT}_1) \) and \( (\alpha_{eq,2}, \text{ROOT}_2) \), with \( \alpha_{eq,1} = \text{ROOT}_1 \) and \( \alpha_{eq,2} = \text{ROOT}_2 \). An illustration is given in Fig. 3.

This completes the proof.

**Fig. 3.** An illustration of two curves \( R_{pp} \) vs. \( \alpha \) and \( R_{ps} \) vs. \( \alpha \) (for a specific \( \rho \)) in Case 1, in which the two curves have up to two intersections.

Below we show some features of \( \text{ROOT}_1 \) and \( \text{ROOT}_2 \). \( \text{ROOT}_1 \) is expressed as

\[
\text{ROOT}_1 = R_{pp}|_{\alpha=\text{ROOT}_1} = \frac{1 - \rho}{2} \log \left[ 1 + \frac{P_g g_{ps} - P_g g_{ps}}{g_{ss}} \frac{g_{ps}}{g_{ss}} - \frac{\sigma^2}{g_{ss} g_{sp}} \right] + L_1 \left( \frac{-L_2 + \sqrt{L_2^2 - 4L_1 L_3}}{2L_1} + \frac{\sigma^2}{g_{ss}} \right). \quad (17)
\]

The second order derivative of \( \text{ROOT}_1 \) is given by

\[
\frac{d^2 \text{ROOT}_1}{d\rho^2} = \frac{1}{(\ln 2)H_1(\rho)} \left( \frac{L_1}{L_2^2 - 4L_1 L_3} \right)^2 \left( \frac{1}{2} \frac{d^2 P_e}{d\rho^2} - \frac{dP_e}{d\rho} \right) - \frac{1 - \rho}{2(\ln 2)^2} \frac{L_1^2}{(L_2^2 - 4L_1 L_3)^2} \left( \frac{dP_e}{d\rho} \right)^2 \quad (18)
\]

\( L_2 \) is not a function of \( \rho \). By including “\( L_2 \geq 0 \)” in the expression of \( \mathcal{A}_4 \), we mean that \( \mathcal{A}_4 \) will be a null set if \( L_2 < 0 \).
where
\[ H_1(\rho) = 1 + \frac{P_y g_{pp}}{\sigma^2} + \frac{P_y g_{ps} - P_y g_{pp}}{g_{ss}} - \frac{\sigma^2}{g_{ss}} - \frac{\sigma^2}{g_{pp}} + L_1 \left( \frac{-L_2 + \sqrt{L_2^2 - 4L_1L_3}}{2L_1} + \frac{\sigma^2}{g_{ss}} \right), \]
and
\[
\begin{align*}
\frac{d^2 P_y}{d \rho^2} &= \frac{2(\rho g_{ps} + E_0 - E_0)}{(1 - \rho)^3} \\
\frac{d^2 P_y}{d \rho^2} &= \frac{4(\rho g_{ps} + E_0 - E_0)}{(1 - \rho)^3}.
\end{align*}
\tag{19}
\]

From (18), since \( \frac{d^2 R_{eq,1}}{d \rho^2} < 0 \), one can obtain that \( \frac{d^2 R_{eq,1}}{d \rho^2} < 0 \), which indicates that \( R_{eq,1} \) is a concave function of \( \rho \).

\[ R_{eq,2} = R_{pp}\big|_{\alpha = \text{root}_2} \]
\[ = \frac{1 - \rho}{2} \log \left[ 1 + \frac{P_y g_{pp}}{\sigma^2} + \frac{P_y g_{ps} - P_y g_{pp}}{g_{ss}^2} + \frac{\sigma^2}{g_{pp}} + L_1 \left( \frac{-L_2 - \sqrt{L_2^2 - 4L_1L_3}}{2L_1} + \frac{\sigma^2}{g_{ss}} \right) \right]. \tag{20} \]

The second order derivative of \( R_{eq,2} \) is given by
\[ \frac{d^2 R_{eq,2}}{d \rho^2} = \frac{-1}{\ln(2)H_2(\rho)} \frac{L_1}{\sqrt{L_2^2 - 4L_1L_3}} \left( 1 - \rho \frac{d^2 P_y}{d \rho^2} - \frac{d P_y}{d \rho} \right) - \frac{1}{\ln(2)(H_2(\rho))^2} \frac{L_1^2}{L_2^2 - 4L_1L_3} \left( \frac{d P_y}{d \rho} \right)^2 \]
\[ + \frac{1}{\ln(2)H_2(\rho)} \left( \frac{L_1^2}{L_2^2 - 4L_1L_3} \right)^2 \frac{d P_y}{d \rho}.
\tag{21} \]

Recall that intersection \((\alpha_{eq,2}, R_{eq,2})\) exists only when \( \rho \in A_5 \). By observing (21), it can be seen that 1) when \( \rho \in A_6 \triangleq A_5 \cap \{ \rho | H_2(\rho) - \frac{1}{2} \sqrt{L_2^2 - 4L_1L_3} \geq 0 \} \), we have \( \frac{d^2 R_{eq,2}}{d \rho^2} \geq 0 \), which indicates that \( R_{eq,2} \) is a convex function of \( \rho \) over \( A_6 \); 2) when \( \rho \in A_7 \triangleq A_5 \cap \{ \rho | H_2(\rho) - \frac{1}{2} \sqrt{L_2^2 - 4L_1L_3} < 0 \} \), we have \( \frac{d^2 R_{eq,2}}{d \rho^2} < 0 \), which indicates that \( R_{eq,2} \) is a concave function of \( \rho \) over \( A_7 \). Note that \( A_4 \) and \( A_7 \) both are closed intervals of \( \rho \).

2) Problem P1 with Case 1: From Lemma 3, when \( \rho \in \bar{A}_4 \), the two curves \( R_{pp} \) vs. \( \alpha \) and \( R_{ps} \) vs. \( \alpha \) do not have intersection, which means that we always have \( R_{pp} < R_{ps} \), as shown in Fig. 3(a). Thus, at optimality of Problem P1, we should have \( R_{pp} = \gamma^T \). Then Problem P1 is equivalent to
\[ \begin{array}{ll}
\max_{\rho} & R_{ss} \big|_{\alpha = \text{F}_p(\rho)} \\
\text{s.t.} & \rho \in A_1 \cap A_2 \cap \bar{A}_4,
\end{array} \tag{22}\]

with \( R_{ss} \big|_{\alpha = \text{F}_p(\rho)} = \frac{1 - \rho}{2} \log \left( 1 + \frac{L_1}{\rho g_{ps} - \mu g_{ss}^2 + P_y g_{pp}} \right). \]

When \( \rho \in A_4 \setminus A_5 \), the two curves have one intersection \((\alpha_{eq,1}, R_{eq,1})\). It is interesting to point out that we further have two scenarios, as follows.

- If \( R_{eq,1} \geq \gamma^T \), then at optimality of Problem P1 (i.e., when \( \min \{ R_{pp}, R_{ps} \} = \gamma^T \)), we have \( R_{ps} \geq R_{pp} \), and thus, we have \( R_{ss} = \gamma^T \) at optimality of Problem P1. Accordingly, Problem P1 is equivalent to
\[ \begin{array}{ll}
\max_{\rho} & R_{ss} \big|_{\alpha = \text{F}_p(\rho)} \\
\text{s.t.} & \rho \in A_1 \cap A_3 \cap \left( \{ A_4 \setminus A_5 \} \cap A_8 \right)
\end{array} \tag{23}\]

in which \( R_{ss} \big|_{\alpha = \text{F}_p(\rho)} = \frac{1 - \rho}{2} \log \left( 1 + \frac{L_1}{\rho g_{ps} - \mu g_{ss}^2 + P_y g_{pp}} \right) \) and \( A_8 \triangleq \{ \rho | R_{eq,1} \geq \gamma^T \} \). As aforementioned, \( R_{eq,1} \) is a concave function of \( \rho \). Thus, \( A_8 \) is a closed interval of \( \rho \).

- If \( R_{eq,1} \leq \gamma^T \), then at optimality of Problem P1 (i.e., when \( \min \{ R_{pp}, R_{ps} \} = \gamma^T \)), we have \( R_{pp} \leq R_{ps} \), and thus, we have \( R_{pp} = \gamma^T \) at optimality of Problem P1. Accordingly, Problem P1 is equivalent to
\[ \begin{array}{ll}
\max_{\rho} & R_{ss} \big|_{\alpha = \text{F}_p(\rho)} \\
\text{s.t.} & \rho \in A_1 \cap A_2 \cap \left( \{ A_4 \setminus A_5 \} \cap A_8 \right)
\end{array} \tag{24} \]

in which \( A_8 \in [0, 1] \setminus A_8 \). As \( A_8 \) is a closed interval, \( A_8 \) is the union of two closed intervals.

When \( \rho \in A_5 \), the two curves have two intersections at \((\alpha_{eq,1}, R_{eq,1})\) and \((\alpha_{eq,2}, R_{eq,2})\). It is also interesting to point out that we further have two scenarios, as follows.

- If \( R_{eq,1} \leq \gamma^T \) or \( R_{eq,2} \geq \gamma^T \), then at optimality of Problem P1 (i.e., when \( \min \{ R_{pp}, R_{ps} \} = \gamma^T \)), we have \( R_{pp} \leq R_{ps} \), and thus, we have \( R_{pp} = \gamma^T \) at optimality of Problem P1. Accordingly, Problem P1 is equivalent to
\[ \begin{array}{ll}
\max_{\rho} & R_{ss} \big|_{\alpha = \text{F}_p(\rho)} \\
\text{s.t.} & \rho \in A_1 \cap A_3 \cap A_5 \cap \left( \bar{A}_8 \cup A_9 \right)
\end{array} \tag{25} \]

in which \( A_9 \in \{ \rho | R_{eq,2} \geq \gamma^T \} \) and \( \rho \in A_9 \cup \{ \rho | R_{eq,2} \leq \gamma^T \} \) and \( \rho \in A_7 \). As \( R_{eq,2} \) is convex over \( A_8 \) and concave over \( A_7 \), \( A_9 \) is the union of three closed intervals.

- If \( R_{eq,1} \geq \gamma^T \) and \( R_{eq,2} \leq \gamma^T \), then at optimality of Problem P1 (i.e., when \( \min \{ R_{pp}, R_{ps} \} = \gamma^T \)), we have \( R_{ps} \leq R_{pp} \), and thus, we have \( R_{ps} = \gamma^T \) at optimality of Problem P1. Accordingly, Problem P1 is equivalent to
\[ \begin{array}{ll}
\max_{\rho} & R_{ss} \big|_{\alpha = \text{F}_p(\rho)} \\
\text{s.t.} & \rho \in A_1 \cap A_3 \cap A_5 \cap A_8 \cap \bar{A}_9
\end{array} \tag{26} \]

in which \( A_9 \in \{ \rho | R_{eq,2} \leq \gamma^T \} \) and \( \rho \in A_9 \cup \{ \rho | R_{eq,2} \leq \gamma^T \} \) and \( \rho \in A_7 \), being the union of three closed intervals.

As a summary, for Case 1, the maximal objective function of Problem P1 is the largest one among the maximal objective functions of Problem (22), Problem (23), Problem (24), Problem (25) and Problem (26).
B. Case 2: When $g_{pp} > g_{ps}$ and $g_{sp} > g_{ss}$

Similar to Lemma 3, when $\rho \in A_4$, the two curves do not have intersection, which means that we always have $R_{ps} \leq R_{pp}$. Thus, at optimality of Problem P1, we should have $R_{ps} = \gamma^T$. Then Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha = F_p(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_2 \cap (A_4 \setminus A_3),$$

(27a)

When $\rho \in A_1 \setminus A_3$, the two curves have one intersection at $(\alpha_{eq1}, R_{eq1})$. We further have two scenarios, as follows.

- If $R_{eq1} \geq \gamma^T$, then at optimality of Problem P1 (i.e., when $\min\{R_{pp}, R_{ps}\} = \gamma^T$), we have $R_{pp} \leq R_{ps}$, and thus, we have $R_{pp} = \gamma^T$ at optimality of Problem P1. Accordingly, Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha = F_p(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_2 \cap [A_4 \setminus A_3].$$

(28a)

- If $R_{eq1} \leq \gamma^T$, then at optimality of Problem P1 (i.e., when $\min\{R_{pp}, R_{ps}\} = \gamma^T$), we have $R_{ps} \leq R_{pp}$, and thus, we have $R_{ps} = \gamma^T$ at optimality of Problem P1. Accordingly, Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha = F_p(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_3 \cap [A_4 \setminus A_5].$$

(29a)

When $\rho \in A_5$, the two curves have two intersections at $(\alpha_{eq1}, R_{eq1})$ and $(\alpha_{eq2}, R_{eq2})$. We further have two scenarios, as follows.

- If $R_{eq1} \leq \gamma^T$ or $R_{eq2} \geq \gamma^T$, then at optimality of Problem P1 (i.e., when $\min\{R_{pp}, R_{ps}\} = \gamma^T$), we have $R_{ps} \leq R_{pp}$ and thus, we have $R_{ps} = \gamma^T$ at optimality of Problem P1. Accordingly, Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha = F_p(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_3 \cap [A_4 \setminus A_5].$$

(30a)

- If $R_{eq1} \geq \gamma^T$ and $R_{eq2} \leq \gamma^T$, then at optimality of Problem P1 (i.e., when $\min\{R_{pp}, R_{ps}\} = \gamma^T$), we have $R_{pp} \leq R_{ps}$ and thus, we have $R_{pp} = \gamma^T$ at optimality of Problem P1. Accordingly, Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha = F_p(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_2 \cap A_5 \cap A_8.$$ 

(31a)

As a summary, for Case 2, the maximal objective function of Problem P1 is the largest one among the maximal objective functions of Problem (27), Problem (28), Problem (29), Problem (30) and Problem (31).

C. Case 3: When $g_{pp} \leq g_{ps}$ and $g_{sp} > g_{ss}$

Similar to the previous two cases, we also try to decide which one between $R_{pp}$ and $R_{ps}$ should be equal to $\gamma^T$.

Lemma 4: For a specific value of $\rho$, the two curves $R_{pp}$ vs. $\alpha$ and $R_{ps}$ vs. $\alpha$ over $\alpha \in [0, 1]$ have at most one intersection at $(\alpha_{eq1}, R_{eq1})$, with $R_{eq1}$ being a concave function of $\rho$.

Proof: The proof is similar to Lemma 3. The major difference is $L_1 > 0$ and $L_2 > 0$, which makes ROOT in (16) larger than 1 and infeasible, yielding at most one intersection.

Note that we have $R_{pp}|_{\alpha = 0} = \frac{1}{2} \log(1 + \frac{g_{pp}}{\alpha}) + \frac{g_{pp}}{\alpha}$; $R_{ps}|_{\alpha = 0} = \frac{1}{2} \log(1 + \frac{g_{ps}}{\alpha}) + \frac{g_{ps}}{\alpha}$. Thus, the two curves $R_{pp}$ vs. $\alpha$ and $R_{ps}$ vs. $\alpha$ have one intersection over $\alpha \in [0, 1]$ if and only if $R_{pp}|_{\alpha = 1} \geq R_{ps}|_{\alpha = 1}$, which equivalently means that $\rho \in A_4 = \left\{ \frac{1}{2} \left( 1 - \frac{g_{pp} \log(1 + \frac{g_{pp}}{\alpha}) + g_{ps}}{g_{ps} + g_{pp}} \right), 0 \right\}$.

Similar to Lemma 3, when $\rho \in A_4 = [0, 1]\setminus A_4$, the two curves do not have intersection, which means that we always have $R_{pp} < R_{ps}$. Thus, at optimality of Problem P1, we should have $R_{pp} = \gamma^T$. Then Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha = F_p(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_2 \cap A_4.$$ 

(32a)

When $\rho \in A_4$, the two curves have one intersection at $(\alpha_{eq1}, R_{eq1})$. We further have two scenarios, as follows.

- If $R_{eq1} \geq \gamma^T$, then at optimality of Problem P1 (i.e., when $\min\{R_{pp}, R_{ps}\} = \gamma^T$), we have $R_{pp} \leq R_{ps}$, and thus, we have $R_{pp} = \gamma^T$ at optimality of Problem P1. Accordingly, Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha = F_p(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_2 \cap [A_4 \setminus A_5].$$ 

(33a)

- If $R_{eq1} < \gamma^T$, then at optimality of Problem P1 (i.e., when $\min\{R_{pp}, R_{ps}\} = \gamma^T$), we have $R_{ps} < R_{pp}$, and thus, we have $R_{ps} = \gamma^T$ at optimality of Problem P1. Accordingly, Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha = F_p(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_3 \cap A_4 \cap [A_8].$$ 

(34a)

As a summary, for Case 3, the maximal objective function of Problem P1 is the largest one among the maximal objective functions of Problem (32), Problem (33), and Problem (34).

D. Case 4: $g_{pp} > g_{ps}$ and $g_{sp} \leq g_{ss}$

Similar to Case 3, we can also prove that Lemma 4 holds in Case 4. Different from Case 3, here we have $R_{pp}|_{\alpha = 0} = \frac{1}{2} \log(1 + \frac{g_{pp}}{\alpha}) + \frac{g_{pp}}{\alpha}$; $R_{ps}|_{\alpha = 0} = \frac{1}{2} \log(1 + \frac{g_{ps}}{\alpha}) + \frac{g_{ps}}{\alpha}$. Thus, the two curves $R_{pp}$ vs. $\alpha$ and $R_{ps}$ vs. $\alpha$ have one intersection over $\alpha \in [0, 1]$ if and only if $R_{pp}|_{\alpha = 1} \leq R_{ps}|_{\alpha = 1}$, which equivalently means that $\rho \in A_4.\text{7}$

When $\rho \in A_4$, the two curves do not have intersection, which means that we always have $R_{pp} > R_{ps}$. Thus, at

\footnote{In Case 3, $R_{pp}|_{\alpha = 1} \geq R_{ps}|_{\alpha = 1}$ is equivalent to $\rho \in A_4$, while in Case 4, $R_{pp}|_{\alpha = 1} \leq R_{ps}|_{\alpha = 1}$ is equivalent to $\rho \in A_4$. This is because we have $g_{pp} \geq g_{ps}$ and $g_{sp} > g_{ss}$ in Case 3, while we have $g_{pp} > g_{ps}$ and $g_{sp} \leq g_{ss}$ in Case 4.}
optimality of Problem P1, we should have $R_{ps} = \gamma^T$. Then Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha=F_\epsilon(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_3 \cap \bar{A}_4.$$  

(35a)

When $\rho \in A_4$, the two curves have one intersection at $(\alpha_{eq1}, R_{eq1})$. We have two scenarios as follows. 

- If $R_{eq1} \geq \gamma^T$, then at optimality of Problem P1 (i.e., when $\min\{R_{pp}, R_{ps}\} = \gamma^T$), we have $R_{pp} \leq R_{pp}$, and thus, we have $R_{ps} = \gamma^T$ at optimality of Problem P1. Accordingly, Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha=F_\epsilon(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_3 \cap A_4 \cap A_8.$$  

(35b)

- If $R_{eq1} < \gamma^T$, then at optimality of Problem P1 (i.e., when $\min\{R_{pp}, R_{ps}\} = \gamma^T$), we have $R_{pp} < R_{ps}$, and thus, we have $R_{pp} = \gamma^T$ at optimality of Problem P1. Accordingly, Problem P1 is equivalent to

$$\max_{\rho} \quad R_{ss}|_{\alpha=F_\epsilon(\rho)} \quad \text{s.t.} \quad \rho \in A_1 \cap A_2 \cap A_4 \cap \bar{A}_8.$$  

(36b)

As a summary, for Case 4, the maximal objective function of Problem P1 is the largest one among the maximal objective functions of Problem (35), Problem (36), and Problem (37).

In Cases 1~4, all the equivalent problems are in the form of $\max_{\rho} R_{ss}|_{\alpha=F_\epsilon(\rho)}$ or $\max_{\rho} R_{ss}|_{\alpha=F_\epsilon(\rho)}$ under a constraint that $\rho$ is within a closed interval (i.e., such as equivalent problems (23), (28), (32) and (35)) or within a union of multiple closed intervals (i.e., such as equivalent problems (25), (30), (34) and (37)). We have two observations:

- If $\rho$ is within a union of multiple closed intervals, we can first get the optimal solution over each interval and pick up the best optimal solution.

- $R_{ss}|_{\alpha=F_\epsilon(\rho)}$ and $R_{ss}|_{\alpha=F_\epsilon(\rho)}$ can be expressed in a unified form as

$$R_{ss,n} = \frac{1}{2} \rho \log \left( 1 + \frac{(P_e g_{sn} - \mu_n \sigma^2 + P_p g_{pn}) g_{ss}}{(\mu_n + 1) g_{sn} \sigma^2} \right)$$  

with $n \in \{p, s\}$. Here we have $R_{ss}|_{\alpha=F_\epsilon(\rho)} = R_{ss,n}$. Therefore, in what follows, we focus on maximizing $R_{ss,n}$, $n \in \{p, s\}$ over $\rho \in \mathcal{B}$ with $\mathcal{B}$ being a closed interval.

$E$. Solving $\max_{\rho} R_{ss,n}$ s.t. $\rho \in \mathcal{B}$

The objective function $R_{ss,n}$ is not a concave function. However, it is a quasiconcave function [27], as theoretically shown in the following theorem.

**Theorem 1**: The objective function $R_{ss,n}$ is quasiconcave with respect to $\rho \in \mathcal{B}$.

**Proof**: See Appendix A.

As shown in [27], super-level sets of a quasiconcave function can be represented by inequalities of concave functions. As reference [27] does not provide methods to find the inequalities of concave functions, here we develop a method to find the inequalities of concave functions that can represent super-level sets of our quasiconcave function $R_{ss,n}$. As a result, we have the following theorem.

**Theorem 2**: For any $t \geq 0$, inequality $R_{ss,n} \geq t$ and inequality $\xi_{t,n} \geq 0$ are equivalent, in which $\xi_{t,n}$ is given as

$$\xi_{t,n} = (\lambda_n - 2\tau \varpi) (1 - \rho) (\mu_n + 1),$$  

(39)

where

$$\lambda_n = 1 + \frac{(P_e g_{sn} - \mu_n \sigma^2 + P_p g_{pn}) g_{ss}}{(\mu_n + 1) g_{sn} \sigma^2}.$$  

(40)

**Proof**: It is readily checked that $R_{ss,n} \geq t$ is equivalent to $\lambda_n - 2\tau \varpi \geq 0$.

Since $1 - \rho$ and $\mu_n + 1$ are larger than 0, multiplying two positive numbers to $\lambda_n - 2\tau \varpi$ will not change its sign. Then we can see that $R_{ss,n} \geq t$ is equivalent to $\xi_{t,n} \geq 0$.

This completes the proof.

From Theorem 2, identifying the maximized $R_{ss,n}$ in $\rho \in \mathcal{B}$ is equivalent to finding the maximal possible value of $t$ such that there exists $\rho \in \mathcal{B}$ that makes $R_{ss,n} \geq t$, which is further equivalent to finding the maximal possible value of $t$ such that there exists $\rho \in \mathcal{B}$ that makes $\xi_{t,n} \geq 0$. For a value of $t$, if there exists $\rho \in \mathcal{B}$ that makes $\xi_{t,n} \geq 0$, then we say that the $t$ value is *feasible*. Thus, max $R_{ss,n}$ s.t. $\rho \in \mathcal{B}$ is equivalent to finding the maximal feasible $t$ value, which can be done by using a bisection search over $t \in [0, \log(1 + \frac{P_e g_{sn}}{\mu_n \sigma^2})]$ (noting that from (4) and (6), it can be seen that an upper bound of $R_{ss,n}$ is $\log(1 + \frac{P_e g_{sn}}{\mu_n \sigma^2})$).

In the bisection search over $t$, we need to decide whether or not a checked $t$ value is feasible. For this purpose, the following theorem, which gives a feature of $\xi_{t,n}$, is helpful.

**Theorem 3**: For a given nonnegative value of $t$, the function $\xi_{t,n}$ is concave with respect to $\rho \in \mathcal{B}$.

**Proof**: The proof is similar to the proof in Appendix C in [1] (conference version of this paper), and thus, is omitted due to space limit.

In the bisection search over $t$, if a checked $t$ value is feasible, this equivalently means that for the checked $t$ value, the maximal value of $\xi_{t,n}$ over $\rho \in \mathcal{B}$ is nonnegative. Thus, from Theorem 3, we can use a bisection search over $\rho \in \mathcal{B}$, to find the maximal value of $\xi_{t,n}$. If any searched $\rho$ makes $\xi_{t,n} \geq 0$, then the checked $t$ value is feasible, and we can terminate the bisection search of $\rho$. If the found maximal value of $\xi_{t,n}$ is negative, then the checked $t$ value is infeasible. Thus, we have two levels of bisection search, and we call the bisection search over $t$ as outer bisection search, and call the bisection search over $\rho$ as inner bisection search.

Generally, each inner bisection search may need to be done over $\rho \in \mathcal{B}$. Next we reduce the number of iterations in inner bisection search.

For a $t$ value, define $\mathcal{R}_{t,n} = \{\rho|\xi_{t,n} \geq 0, \rho \in \mathcal{B}\}$. So $\mathcal{R}_{t,n}$ is a set of $\rho$ that makes $\xi_{t,n} \geq 0$. Consider two $t$ values: $t_1$ and $t_2$, with $t_1 < t_2$. We have

$$\mathcal{R}_{t_2,n} \supseteq \mathcal{R}_{t_1,n} = \{\rho|R_{ss,n} \geq t_1\} \supseteq \{\rho|R_{ss,n} \geq t_2\}.$$  

(41)

In the bisection search, we try to find a value of $\rho$ that makes $\xi_{t,n} = 0$ (which means the maximal value of $\xi_{t,n}$ is achieved over $\rho \in \mathcal{B}$). Note that other methods can also be used here to find the maximal value of $\xi_{t,n}$ over $\rho \in \mathcal{B}$, such as gradient descent method and Newton’s method.
in which steps (i) and (iii) are from Theorem 2, and step (ii) is from $t_1 < t_2$.

From (41), we can see that, if we know that $t_1$ is feasible, then when deciding whether or not $t_2$ is feasible, we only need to search $\rho$ over $R_{t_1,n}$, or over an interval of $\rho$ that includes $R_{t_1,n}$ as a subset. Denote $\mathcal{F} \triangleq [f_1, f_2]$ as the interval of $\rho$ over which the inner bisection search is performed. So $\mathcal{F}$ is initially set to $B$. In the outer bisection search, for a checked $t$ value (say $t^1$), if $t^1$ is feasible, we get an updated $\mathcal{F}$ (which is a subset of the previous $\mathcal{F}$, and includes $R_{t_1,n}$ as a subset). Then, in the outer bisection search, when we check feasibility of higher $t$ values, we only need to search $\rho$ over the updated interval $\mathcal{F}$ (rather than over $B$) in the inner bisection search, referred to as feasible region shrinking.

Based on this observation, we have Algorithm 1 for our inner bisection search. In Algorithm 1, the expression of $\xi_{t,n}$ is given in (39), while expression of $\frac{d\xi_{t,n}}{d\rho}$ is given as

$$
\frac{d\xi_{t,n}}{d\rho} = (g_{ss} - 1)(\mu_n - (1 - \rho)\frac{d\mu_n}{d\rho}) + \frac{2n}{\sigma^2}P_g g_{ss} \frac{d\xi_{t,n}}{d\rho} - (g_{ss} + 1)(2\frac{(\ln 2)t}{1 - \rho} - 1)(1 - \rho)\frac{d\mu_n}{d\rho})2\frac{\Delta}{\sigma^2}.
$$

(42)

**Algorithm 1: Inner Bisection Search**

**Input:** value of $t$ to be checked for feasibility, and two end points of the interval $\mathcal{F}$: $f_1$ and $f_2$

**Output:** Feasibility of value $t$, updated $f_1$ and $f_2$

1. if $\xi_{t,n}|_{\rho=f_1} \geq 0$ or $\xi_{t,n}|_{\rho=f_2} \geq 0$ then
2. $t$ is infeasible. Terminate the algorithm;
3. if $\frac{d\xi_{t,n}}{d\rho}|_{\rho=f_1} \times \frac{d\xi_{t,n}}{d\rho}|_{\rho=f_2} \geq 0$ then
4. $t$ is infeasible. Terminate the algorithm;
5. $f_1^{BS} \leftarrow f_1, f_2^{BS} \leftarrow f_2$;
6. $f_{new} \leftarrow f_1^{new}, f_{new}^{new} \leftarrow f_2$;
7. if $|f_{new}^{BS} - f_1^{BS}| < \epsilon_{\rho}$ then
8. $t$ is infeasible. Terminate the algorithm;
9. $f_{mid} \leftarrow (f_1^{BS} + f_2^{BS})/2$;
10. if $\xi_{t,n}|_{\rho=f_{mid}} \geq 0$ then
11. $t$ is feasible; $f_1 \leftarrow f_{new}^{new}, f_2 \leftarrow f_2^{new}$; Terminate the algorithm;
12. if $\frac{d\xi_{t,n}}{d\rho}|_{\rho=f_{mid}} = 0$ then
13. $t$ is infeasible. Terminate the algorithm;
14. if $\frac{d\xi_{t,n}}{d\rho}|_{\rho=f_{mid}} > 0$ then
15. $f_1^{BS} \leftarrow f_{mid}, f_2^{new} \leftarrow f_{mid}$; Go to Step 7;
16. if $\frac{d\xi_{t,n}}{d\rho}|_{\rho=f_{mid}} < 0$ then
17. $f_2^{BS} \leftarrow f_{mid}, f_{new} \leftarrow f_{mid}$; Go to Step 7;

In Steps 1 and 2 of Algorithm 1, we check whether $\rho = f_1$ or $\rho = f_2$ makes $\xi_{t,n}$ nonnegative. If yes, then the checked $t$ value is feasible and we terminate the algorithm. If we proceed to Step 3, then we know that $\xi_{t,n}|_{\rho=f_1} < 0$ and $\xi_{t,n}|_{\rho=f_2} < 0$. In Step 3, we check whether $\frac{d\xi_{t,n}}{d\rho}|_{\rho=f_1}$ and $\frac{d\xi_{t,n}}{d\rho}|_{\rho=f_2}$ are both nonpositive or both nonnegative. If they are both nonpositive or both nonnegative, then $\xi_{t,n}$ is a decreasing or increasing function with respect to $\rho \in \mathcal{F}$, and thus, we can conclude that the checked $t$ value is infeasible. So when we proceed to Step 5, we should have $\frac{d\xi_{t,n}}{d\rho}|_{\rho=f_1} > 0$ and $\frac{d\xi_{t,n}}{d\rho}|_{\rho=f_2} < 0$. Then we try to search the maximal point of $\xi_{t,n}$ over $\rho \in [f_1, f_2]$, by using inner bisection search of $\frac{d\xi_{t,n}}{d\rho}$ until $\frac{d\xi_{t,n}}{d\rho} = 0$. In the inner bisection search, $[f_1^{BS}, f_2^{BS}]$ represents the subinterval after bisecting the original interval of $\rho$, and $[f_1^{new}, f_2^{new}]$ represents the updated interval $\mathcal{F}$. In the inner bisection search, if a searched $\rho$ value makes $\xi_{t,n} \geq 0$, then we know that the checked $t$ value is feasible, and we update $\mathcal{F}$ (we can see that the updated $\mathcal{F}$ satisfies $R_{t,n} \subseteq \mathcal{F}$), and terminate the algorithm (Steps 10–11). If the subinterval $[f_1^{BS}, f_2^{BS}]$ is sufficiently small (i.e., less than a threshold value $\epsilon_{\rho}$) and no searched $\rho$ value makes $\xi_{t,n} \geq 0$, then we know the checked $t$ value is infeasible, and we do not update $\mathcal{F}$ (Steps 7–8).

The detailed algorithm for the outer bisection search is straightforward, and thus, is omitted here, for presentational simplicity.

**Complexity:** According to [28], the computational complexity of a bisection search is $O(\log(\epsilon))$, where $\epsilon$ is the pre-defined tolerance for convergence. Thus, the complexity of the proposed two-level bisection search is expressed as $O(\log(\epsilon))$, in which $\epsilon_{t}$ and $\epsilon_{\rho}$ are pre-defined convergence tolerance for the outer bisection search over $t$ and the inner bisection search over $\rho$.

**Impact of parameters $\epsilon_{t}$ and $\epsilon_{\rho}$:** Denote $t^*$ as the maximal feasible $t$ value, and denote $\rho^*$ as the corresponding $\rho$ that achieves $t^*$. Denote $\hat{t}$ as the $t$ value found by our proposed algorithm.

For the outer bisection search over $t$, when it converges, we get a region of $t$, denoted as $[l, u]$ with $u - l < \epsilon_t$. Then $\hat{t} = l$ is the $t$ value found by our algorithm.

If the inner bisection search (which does feasibility check for a specific $t$ value) is always accurate, then when the outer bisection search converges, $t = l$ is feasible, $t = u$ is infeasible, and $t^*$ falls within $[l, u]$. Thus, the gap between $t^*$ and $\hat{t}(= l)$ is less than $\epsilon_t$.

However, when the inner bisection search checks the feasibility of a value, say $t^1$, close to $t^*$, it may not be accurate. It is possible that the inner bisection search may claim that $t^1$ is infeasible, but actually $t^1$ is feasible. If this happens, then when the outer bisection search converges at a region of $t$ denoted as $[l, u]$, actually $t = l$ and $t = u$ are both feasible, and we have $l < u < t^*$ (in other words, $t^*$ does not fall within $[l, u]$). Recall that our objective function $R_{ss,n}$ is a quasiconcave function, i.e., when the first order derivative is 0, the second order derivative is nonpositive. Thus, we can treat the quasiconcave function $R_{ss,n}$ as a concave function at the neighborhood of $\rho^*$. Based on this, $|t^* - u|$ is less than $\frac{|\epsilon_{t} - \epsilon_{\rho}|}{|dR_{ss,n}|_{\rho=\rho^*}}$. Accordingly, the gap between $t^*$ and $\hat{t}(= l)$ is less than $\Delta \triangleq \epsilon_t + \frac{|\epsilon_{t} - \epsilon_{\rho}|}{|dR_{ss,n}|_{\rho=\rho^*}}\epsilon_{\rho}$.

For $R_{ss,n}$, its first-order derivative at $\rho = \rho^*$ is zero, i.e.,
Since $R_{ss,n}$ can be viewed as a concave function at the neighborhood of $\rho^*$ and $\epsilon_\rho$ is small, we can see that $\left.\frac{dR_{ss,n}}{d\rho}\right|_{\rho=\rho^* - \epsilon_\rho}$ is close to zero. Thus, $\Delta$ is small for small values of $\epsilon_t$ and $\epsilon_\rho$.

F. Further Discussion: When Primary Information is Sent during the First Phase

In the first phase, the wireless RF signals can actually be information signals of the PT for the PR. In other words, in the first phase, from the PT’s signals, the ST tries to harvest energy while the PR tries to decode information. Thus, the primary system has an additional throughput expressed as $\rho \log(1 + \frac{P_p g_{pp}}{\sigma^2})$. It also means that the target throughput of the primary system during the second and third phase is $\gamma^T - \rho \log(1 + \frac{P_p g_{pp}}{\sigma^2})$. Therefore, in Problem P1, we should replace $\gamma^T$ with $\gamma^T - \rho \log(1 + \frac{P_p g_{pp}}{\sigma^2})$, and accordingly, we have the following revised Problem P1:

Revised Problem P1:

$$\max_{\alpha, \rho} R_{ss} = \frac{1 - \rho}{2} \log \left( 1 + \frac{(1 - \alpha) g_{ss}}{\sigma^2} \right)$$  \hspace{1cm} (43a)

s.t.

$$\rho \in \tilde{A}_1; 0 \leq \alpha \leq 1$$

$$\min \{ R_{pp}, R_{ps} \} \geq \gamma^T - \rho \log(1 + \frac{P_p g_{pp}}{\sigma^2})$$  \hspace{1cm} (43c)

in which

$$\tilde{A}_1 \triangleq \left[ \max \left\{ \frac{E_n}{\eta T_p g_{tn}}(0), 1 - \frac{2(\gamma^T - \log(1 + \frac{P_p g_{pp}}{\sigma^2}))}{\log(1 + \frac{P_p g_{pp}}{\sigma^2}) - 2 \log(1 + \frac{P_p g_{pp}}{\sigma^2})} \right\} \right].$$

Similar to Lemma 1, we should have either $R_{pp} = \gamma^T - \rho \log(1 + \frac{P_p g_{pp}}{\sigma^2})$ or $R_{ss} = \gamma^T - \rho \log(1 + \frac{P_p g_{pp}}{\sigma^2})$ at optimality of the revised Problem P1. Accordingly, when $R_{pn} = \gamma^T - \rho \log(1 + \frac{P_p g_{pp}}{\sigma^2})$, $n \in \{p, s\}$, we can see that $\alpha$ can be expressed in terms of $\rho$ as $\alpha = \tilde{F}_n(\rho) \triangleq \frac{\tilde{\mu}_n(P_e g_{sn} + \sigma^2) - P_p g_{pp}}{P_p g_{pp} + \tilde{\mu}_n(P_e g_{sn} + \sigma^2)}$ with $\tilde{\mu}_n \triangleq \frac{(1 + \frac{P_p g_{pp}}{\sigma^2})^2 \sigma^2}{2 \sigma^2 - \frac{P_p g_{pp}}{\sigma^2} - 1}$, in which $\gamma^T = \gamma^T - \log(1 + \frac{P_p g_{pp}}{\sigma^2})$.

Afterwards, for each of the four cases defined in Section III-A~III-D, the revised Problem P1 is equivalent to maximizing

$$R_{ss}|_{\alpha = \tilde{F}_n(\rho)} = \frac{1 - \rho}{2} \log \left( 1 + \frac{(P_e g_{sn} - \tilde{\mu}_n \sigma^2 + P_p g_{pp})g_{sn}}{\tilde{\mu}_n + 1} g_{sn} \sigma^2 \right)$$  \hspace{1cm} (44)

over closed intervals.

Then similar to Theorem 1, we can still prove that $R_{ss}|_{\alpha = \tilde{F}_n(\rho)}$ is quasiconcave over each closed interval, and thus, a two-level bisection search can be used to find the optimal solution.

For presentation simplicity, in the sequel, we consider that the PT does not send information to the PR during the first phase of each slot in the cooperative transmission mode.

IV. OPTIMAL SOLUTION FOR PROBLEM P2

Similar to Problem P1, when Problem P2 achieves the optimality, the constraint $R_{pp} \geq \gamma^T$ should take equality,

\begin{equation}
\frac{dR_{ss,n}}{d\rho}|_{\rho=\rho^*} = 0.
\end{equation}

which means $\alpha = F_p(\rho)$, as given in (9). Accordingly, Problem P2 is equivalent to

$$\max_{\rho} R_{ss}^{w/o}|_{\alpha = F_p(\rho)}$$  \hspace{1cm} (45a)

s.t.

$$\rho \in \tilde{A}_1 \cap \tilde{A}_2$$

with $R_{ss}^{w/o}|_{\alpha = F_p(\rho)} = \frac{1 - \rho}{2} \log \lambda^{w/o}$, where

$$\lambda^{w/o} = 1 + (P_e g_{sp} - \mu_p \sigma^2 + P_p g_{pp}) g_{ss} \left[ \frac{\mu_p (P_e g_{sp} + \sigma^2) - P_p g_{pp}}{P_p g_{pp}} \right] g_{ss} + (P_p g_{ps} + \sigma^2)(\mu_p + 1) g_{sp} \right]^{-1}}. $$

We have the following theorem for the objective function of Problem (45).

Theorem 4: The objective function $R_{ss}^{w/o}|_{\alpha = F_p(\rho)}$ is quasiconcave with respect to $\rho \in [0, 1]$.

\textbf{Proof:} See Appendix B.

Since $R_{ss}^{w/o}|_{\alpha = F_p(\rho)}$ is quasiconcave, similar to Theorem 2, for any $t \geq 0$, inequality $R_{ss}^{w/o}|_{\alpha = F_p(\rho)} \geq t$ is equivalent to $\xi_t^{w/o} \geq 0$, with

$$\xi_t^{w/o} \triangleq (\lambda^{w/o} - 2 - \frac{2\rho}{1 - \rho}) (1 - \rho) \left[ \frac{\mu_p (P_e g_{sp} + \sigma^2) - P_p g_{pp}}{P_p g_{pp}} \right] g_{ss} + (P_p g_{ps} + \sigma^2)(\mu_p + 1) g_{sp} \right]^{-1}}. $$

The following theorem gives a feature of $\xi_t^{w/o}$.

Theorem 5: For a given nonnegative value $t$, the function $\xi_t^{w/o}$ is concave with respect to $\rho \in [0, 1]$.

\textbf{Proof:} See Appendix C.

Therefore, a two-level bisection search similar to that in Section III-E can be used to find the optimal solution for Problem P2. The details are omitted here.

Overall, Fig. 4 shows the procedure for finding optimal solution of the considered system.

\textbf{Remark:} Interestingly, we have an observation that the maximal secondary throughput of Problem P1 may not guarantee to be larger than that of Problem P2. The reason is as follows. For specific $\alpha$ and $\rho$, the objective function value of Problem P1 is indeed larger than that of Problem P2. However, compared to Problem P2, Problem P1 has one more constraint $R_{ps} \geq \gamma^T$, which makes the feasible region of Problem P1
be a subset of the feasible region of Problem P2. Thus, in a larger feasible region, it is possible that the maximum secondary throughput of Problem P2 is larger than that of Problem P1.

A numerical example is also given here, which has the following parameter setting as: \( g_{tt} = 7.15 \times 10^{-3}, g_{pp} = 2.69 \times 10^{-5}, g_{ps} = 2.65 \times 10^{-6}, g_{sp} = 1.1 \times 10^{-1}, g_{ss} = 2.88 \times 10^{-2}, P_p = 15 \) dBm, \( E_0 = 3.41 \times 10^{-5} \) J, \( E_c = 1.5 \times 10^{-6} \) J, \( \sigma^2 = -25 \) dBm, and \( \gamma^T = 0.5 \) bps/Hz.

In this specific example, the optimal solution of Problem P2 is 0.1448 bps/Hz, which is larger than the optimal solution of Problem P1, 0.0321 bps/Hz.

V. Numerical Results

We use Matlab simulation to evaluate the performance of our proposed algorithm. Similar to [16], the channel power gain \( g_i \) is further represented as \( g_i = \frac{\gamma_i}{1 + d_i^\alpha} \), where \( i \in \mathcal{M} = \{tt, pp, ps, sp, ss\} \), \( \gamma_i \) is exponentially distributed with parameter 1, and \( d_i \) is the distance of link \( i \), and \( \alpha \) is the path loss exponent of link \( i \). The distance values are: \( d_{pp} = 12 \) m, \( d_{tt} = d_{ps} = d_{sp} = d_{ss} = 8 \) m. The background noise variance is \( \sigma^2 = -25 \) dBm and the energy conversion efficiency is \( \eta = 0.5 \). The energy expenditure in a time slot due to circuit operation and channel estimation is \( E_c = 1.5 \times 10^{-6} \) J. In our simulation, a minimum throughput requirement 0.15 bps/Hz is set up at the ST. If the ST cannot achieve this throughput at a slot, our system works in direct primary transmission mode.

First we evaluate the chance of ST to access the channel. Three events, \( D_1, D_2 \) and \( D_3 \), are considered. Specifically, event \( D_1 \) is defined as \( g_{pp} < (2^{\gamma^T} - 1)\sigma^2/P_p \), which means that the PT needs help. Event \( D_2 \) is defined as \( g_{tt} > (2^{\gamma^T} - 1)\sigma^2/P_p \), which means that it is possible to set up a \( \rho \) such that \( R_{tt} \geq \gamma^T \) (i.e., the ST is able to decode information from the PT in the second phase). Event \( D_3 \) is defined as the event that the PT needs help (i.e., \( g_{pp} < (2^{\gamma^T} - 1)\sigma^2/P_p \)) and at least one of Problems P1 and P2 is feasible. Clearly, when event \( D_3 \) happens, events \( D_1 \) and \( D_2 \) should also happen. And the probability of event \( D_3 \) is exactly the probability of the ST to access the channel. By setting \( \gamma^T = 0.25 \) bps/Hz and 0.5 bps/Hz, \( \sigma^2 = 4, \alpha = 2, i \in \mathcal{M}\backslash\{pp\} \) and varying \( P_p \) from 0 dBm to 30 dBm, the occurrence probabilities of three events, \( P(D_1), P(D_2), \) and \( P(D_3) \), are shown in Fig. 5. Probabilities \( P(D_1) \) and \( P(D_2) \) are decreasing and increasing functions, respectively, of \( P_p \), which is intuitive. Moreover, we get higher \( P(D_1) \) and lower \( P(D_2) \) through increasing \( \gamma^T \) from 0.25 bps/Hz to 0.5 bps/Hz. This is because a higher target rate \( \gamma^T \) means that the PT needs more help and the ST needs better channel gain \( g_{tt} \) to decode information from the PT. At low \( P_p \), \( P(D_1) \) is close to 1 (i.e., the PT almost always needs help from the ST), and thus, the curve of \( P(D_3) \) follows the trend of the curve of \( P(D_1) \), i.e., increases when \( P_p \) increases or \( \gamma^T \) decreases. At high \( P_p \), \( P(D_2) \) is close to 1, and thus, the curve of \( P(D_3) \) follows the trend of the curve of \( P(D_1) \), i.e., decreases when \( P_p \) increases or \( \gamma^T \) decreases.

Next we show how much chance the performance of Problem P2 is better than that of Problem P1 (i.e., the maximal secondary throughput without SIC is higher than that with SIC). We set \( \gamma^T = 0.1, 0.25, 0.5, 0.75 \), and 1 bps/Hz, and \( P_p = 0 \) dBm and 10 dBm. For each \( (\gamma^T, P_p) \) pair, we run simulations for \( 10^6 \) time slots. In Table I, the numbers after slash are numbers of time slots in which both Problems P1 and P2 are feasible, while the numbers before slash are numbers of time slots when the maximal secondary throughput of Problem P2 is larger than that of Problem P1. It can be seen that, in most time slots when both problems are feasible, Problem P1 has better performance. However, it is still possible (with a small probability) that Problem P2 performs better. Therefore, when both problems are feasible, if we directly take the solution of Problem P1, we have a very large chance to get the overall optimal solution, to be verified below.

We compare our proposed algorithm with its two variants: an SIC-if-possible algorithm that takes the solution of Problem P1 when both Problem P1 and Problem P2 are feasible, and a Never-SIC algorithm that always takes the solution of Problem P2 even if both Problem P1 and Problem P2 are feasible. Fig. 6 shows secondary throughput performance of our proposed algorithm and the two variants for \( \gamma^T = 0.25 \) bps/Hz and 0.5 bps/Hz. It can be seen that the throughput of our proposed algorithm has similar trend as the channel access probability (the curve of \( P(D_3) \) as shown in Fig. 5). By comparing Fig. 5 and Fig. 6, when \( P_p \) increases beyond 20 dBm, the channel access probability and throughput of the secondary system in the proposed algorithm both decrease, but the decrease rate of the channel access probability is higher, explained as follows. When \( P_p \) is high, the channel access probability is low, which means the ST has more chance to accumulate energy. Thus, at a slot, when the system works in cooperative transmission mode, the energy level of the ST is high, leading to high throughput.

\(^{10}\)In the sequel, by “our proposed algorithm,” we mean the algorithm that takes the better solution of Problem P1 and Problem P2 if both problems are feasible, and takes the solution of Problem P2 if only Problem P2 is feasible. Note that if Problem P2 is infeasible, then Problem P1 should also be infeasible.

\(^{11}\)Note that in cooperative transmission mode, when Problem P1 is feasible, Problem P2 is always feasible. Thus, “at least one of Problems P1 and P2 is feasible” is equivalent to “Problem P2 is feasible.”
at the slot. Thus, when $P_p$ increases beyond 20 dBm, the decrease rate of secondary system throughput is not as high as that of the channel access probability. From Fig. 6, the SIC-if-possible algorithm achieves almost the same performance as that of our proposed algorithm. This verifies our statement that when both Problem P1 and Problem P2 are feasible, if we directly take the solution of Problem P1, we have a very large chance to get the overall optimal solution. From Fig. 6, it can also be seen that there is a gap between the performance of the Never-SIC algorithm and that of our proposed algorithm. This is because it is with a very large probability (for example, more than 90% based on simulation results in Table I) that the case with SIC performs better than the case without SIC.

**Remark:** The above observations do not mean that our Problem P2 is useless. This is because it is possible that Problem P1 is infeasible but Problem P2 is feasible, in which scenario we have to take the solution of Problem P2.

In our proposed algorithm, we use two levels of bisection search. In the inner bisection search, we keep shrinking the feasible region as shown in Algorithm 1. To show the benefit of this, we compare two scenarios: our proposed algorithm is implemented with and without shrinking the feasible region in inner bisection search. Recall that the bisection in the inner bisection search may not be implemented in Algorithm 1 (for example, when $\xi_{t,n}|_{p_1} \geq 0$ or $\xi_{t,n}|_{p_2} \geq 0$, Algorithm 1 terminates at Step 2, and thus, the bisection is not implemented). Thus, in the comparison, we only consider the time slots when the bisection in the inner bisection search is carried out (e.g., Algorithm 1 proceeds to Step 5). The average number of inner-bisecting iterations in a two-level bisection search algorithm is shown in Fig. 7, with the tolerance of convergence for bisection search being $\epsilon_1 = \epsilon_0 = 10^{-4}$ and $\epsilon_1 = \epsilon_p = 10^{-4}$. It can be seen that shrinking feasible region in inner bisection search reduces the number of iterations in the inner bisection search by around 10%.

Now we try to compare our proposed algorithm with other algorithms. Since no existing work in the literature considers overlay cognitive NOMA enhanced with TS-based SWIPT, here we compare with two algorithms: an OMA algorithm that partitions one time slot into one harvesting phase and three equal-length information transmission phases for links PT$\rightarrow$ST&PR, ST$\rightarrow$PR, and ST$\rightarrow$SR, respectively, and an equal power allocation (EPA) algorithm that allocates equal amount of energy for transmitting the PT’s signal and the ST’s signal in the third phase of a time slot. Fig. 8 shows secondary throughput of our proposed algorithm, and the OMA and EPA algorithms. Clearly, our proposed algorithm outperforms the EPA and OMA algorithms in terms of higher secondary throughput. Moreover, performance gap of our algorithm with the EPA and OMA algorithms shrink at high $P_p$, explained as follows. At high $P_p$, the probability that the PT needs help is low (as observed in Fig. 5). So at time slots when the PT does not need help, the ST will only accumulate energy. Thus, at a time slot when the PT needs help, the ST has a large chance to have high energy, leading to high signal-to-noise ratio (SNR) at the SR. The throughput is a logarithm function of the SNR. The logarithm function is a concave function, i.e., at high SNR, the probability that the PT needs help is high (as observed in Fig. 5). So at time slots when the PT does not need help, the ST will only accumulate energy. Thus, at a time slot when the PT needs help, the ST has a large chance to have high energy, leading to high signal-to-noise ratio (SNR) at the SR. The throughput is a logarithm function of the SNR. The logarithm function is a concave function, i.e., at high SNR, increase of SNR does not lead to much increase in throughput. Thus, the throughput gap of our algorithm with the EPA and OMA algorithms shrink at high $P_p$.

Distance plays a critical role in SWIPT-based systems, since

### Table I

<table>
<thead>
<tr>
<th>$P_p$ = 0 dBm</th>
<th>$P_p$ = 10 dBm</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma^T = 0.1$</td>
<td>$134/111690$</td>
</tr>
<tr>
<td>$\gamma^T = 0.25$</td>
<td>$150/74603$</td>
</tr>
<tr>
<td>$\gamma^T = 0.5$</td>
<td>$57/46421$</td>
</tr>
<tr>
<td>$\gamma^T = 0.75$</td>
<td>$5/31032$</td>
</tr>
<tr>
<td>$\gamma^T = 1$</td>
<td>$0/21085$</td>
</tr>
</tbody>
</table>

**Fig. 6.** Secondary throughput of the proposed algorithm, the SIC-if-possible algorithm, and the Never-SIC algorithm.

**Fig. 7.** The average number of inner-bisecting iterations with and without shrinking.
Fig. 9. The impact of distance on secondary throughput of our proposed algorithm, the OMA and the EPA algorithms.

The distance between the PT and ST has a big effect on the secondary throughput.

APPENDIX A
THE PROOF OF THEOREM 1

Define \( \lambda_n = 1 + \frac{P_g g_{ss} - \mu_n \sigma^2 + P_g g_{sn} \sigma^2}{(\mu_n + 1) g_{ss} \sigma^2} \). The first and second order derivatives of objective function \( R_{ss,n} \) are derived as

\[
\begin{align*}
\frac{dR_{ss,n}}{dp} & = -\frac{1}{2} \log(\lambda_n) + \frac{1}{2} \frac{1-\rho}{\lambda_n} \frac{d\lambda_n}{dp} + \frac{1}{2} \frac{1-\rho}{\lambda_n} \frac{d\lambda_n}{dp}, \\
\frac{d^2 R_{ss,n}}{dp^2} & = -\frac{1}{2} \frac{1-\rho}{\lambda_n} \frac{d^2 \lambda_n}{dp^2} + \frac{1}{2} \frac{1-\rho}{\lambda_n} \frac{d^2 \lambda_n}{dp^2} + \frac{1}{\lambda_n} \left( \frac{d^2 \lambda_n}{dp^2} \right)^2 + \frac{d^2 \lambda_n}{dp^2},
\end{align*}
\]

(48)

The expression of \( \lambda_n \) can be rewritten as \( \lambda_n = 1 + \frac{\lambda}{\lambda_n} \), where \( \lambda_n \equiv (P_g g_{ss} - \mu_n \sigma^2 + P_g g_{sn} \sigma^2)/(\mu_n + 1) g_{ss} \sigma^2 \). Then the first and second order derivatives of \( \lambda_n \) are given by

\[
\begin{align*}
\frac{d\lambda_n}{dp} & = \frac{1}{\lambda_n} \frac{d\lambda_n}{dp} - \frac{\lambda}{\lambda_n} \frac{d\lambda_n}{dp}, \\
\frac{d^2 \lambda_n}{dp^2} & = \frac{2}{\lambda_n} \frac{d^2 \lambda_n}{dp^2} + \lambda \frac{d^2 \lambda_n}{dp^2} - \lambda \frac{d^2 \lambda_n}{dp^2}.
\end{align*}
\]

(49)

with \( \frac{d^2 P_g}{dp^2} \) and \( \frac{d^2 P_g}{dp^2} \) given in (19).

From (4), an equivalent form of \( \lambda_n \) is \( \lambda_n = 1 + \frac{(1-\rho) P_g g_{ss}}{\lambda_n} \), which is always not less than 1. Since \( \lambda_n = 1 + \frac{\lambda}{\lambda_n} \), we always have \( \mu_n > 0 \), we have \( \lambda_n \geq 0 \).
According to [27, page 101], for a function over an interval, if its second order derivative is nonpositive when its first order derivative is 0, then the function is quasiconcave over the interval. Thus, next, we will show $R_{ss,n}$ is strictly quasiconcave, that is, $\frac{d^2R_{ss,n}}{d\rho^2} < 0$ when $\frac{dR_{ss,n}}{d\rho} = 0$. So in the rest of this proof, we only consider what happens when $\frac{dR_{ss,n}}{d\rho} = 0$.

Based on the first equation of (48), $\frac{dR_{ss,n}}{d\rho} = 0$ is equivalent to

$$-\lambda_n \log(\lambda_n) + \frac{1 - \rho}{\ln(2)} \frac{d\lambda_n}{d\rho} = 0. \tag{51}$$

As $\lambda_n$ is always not less than 1, $\lambda_n \log(\lambda_n)$ is nonnegative. Together with (51), we have $\frac{d\lambda_n}{d\rho} \geq 0$. From the second equation of (49), we have $\frac{d^2\lambda_n}{d\rho^2} = \frac{-2d\lambda_n}{\lambda_n^2} + \omega \leq \omega$, where $\omega = \frac{\lambda_n^2\lambda_n^2 - \lambda_n^2}{\lambda_n^2}$. Here the inequality is because $\frac{d\lambda_n}{d\rho} > 0$ (from (50)), $\frac{d\lambda_n}{d\rho} \geq 0$, and $\lambda_b > 0$ (from definition of $\lambda_b$).

Using $\frac{d^2\lambda_n}{d\rho^2} \leq \omega$, we can obtain the following for the second order derivative of $R_{ss,n}$

$$\frac{d^2R_{ss,n}}{d\rho^2} = \frac{1}{2(\ln(2))\lambda_n^2} + \frac{1}{2(\ln(2))\lambda_n} \tag{52}$$

in which the first equality is from (48). Apparently, for the right hand side of the inequality in (52), the first term is nonpositive. Next we show that the second term is negative. The numerator of the second term can be rewritten as

$$(1 - \rho)\omega - 2\frac{d\lambda_n}{d\rho} = \frac{(1 - \rho)\frac{d^2\lambda_n}{d\rho^2} - 2\frac{d\lambda_n}{d\rho}}{(1 - \rho)\frac{d^2\lambda_n}{d\rho^2} - 2\frac{d\lambda_n}{d\rho} \lambda_n^2 - \frac{(1 - \rho)\frac{d^2\lambda_n}{d\rho^2} - 2\frac{d\lambda_n}{d\rho}}{(1 - \rho)\frac{d^2\lambda_n}{d\rho^2} - 2\frac{d\lambda_n}{d\rho} \lambda_n^2}. \tag{53}$$

For the two terms on the right-hand-side of (53), we have

$$\begin{align*}
(1 - \rho)\frac{d^2\lambda_n}{d\rho^2} - 2\frac{d\lambda_n}{d\rho} & = (1 - \rho)(g_{ssn} \frac{d^2\mu_n}{d\rho^2} - \sigma^2 \frac{d\mu_n}{d\rho} \sigma^2)g_{ss} \\
& - 2(g_{ssn} \frac{d^2\mu_n}{d\rho^2} - \sigma^2 \frac{d\mu_n}{d\rho} \sigma^2)g_{ss} \\
& = -(1 - \rho)\frac{d^2\mu_n}{d\rho^2} - 2\frac{d\mu_n}{d\rho} \sigma^2 g_{ss} \quad \text{iv} < 0 \\
(1 - \rho)\frac{d^2\mu_n}{d\rho^2} - 2\frac{d\mu_n}{d\rho} & = ((1 - \rho)\frac{d\mu_n}{d\rho} - 2\frac{d\mu_n}{d\rho})g_{ss} \sigma^2 \quad \text{v} > 0,
\end{align*} \tag{54}$$

in which steps (iv) and (v) come from the fact that

$$(1 - \rho)\frac{d^2\mu_n}{d\rho^2} - 2\frac{d\mu_n}{d\rho} = \frac{2(\ln(2))\lambda_n^2}{(1 - \rho)^3} 2^{\frac{2}{\lambda_n}} > 0 \tag{55}$$

with step (vi) from (50).

Thus, the numerator of the second term on the right hand-side of the inequality in (52) is negative, and accordingly, we have $\frac{d^2R_{ss,n}}{d\rho^2} < 0$ when $\frac{dR_{ss,n}}{d\rho} = 0$. Therefore, $R_{ss,n}$ is a quasiconcave function. This completes the proof.

**APPENDIX B**

**THE PROOF OF THEOREM 4**

Denote $\lambda^{w/o} = 1 + \frac{\lambda_a - \lambda_{1,\rho}}{\lambda_a}$, where $\lambda_a |_{\rho=1}$ and its first and second order derivatives are given in Appendix A, and $\lambda_c$ and its first and second order derivatives are expressed as

$$\begin{align*}
\lambda_c &= [(\mu_p P_{cgs} + \sigma^2) - P_{pgps}] g_{ss} + (P_p g_{ps} + \sigma^2)(\mu_p + 1) g_{ss}, \\
\frac{d\lambda_c}{d\rho} &= (P_p g_{ps} + \sigma^2) g_{ss} \frac{d\mu_p}{d\rho} + \mu_p g_{ps} \frac{d\mu_p}{d\rho} P_{sp}, \\
\frac{d^2\lambda_c}{d\rho^2} &= (P_p g_{ps} + \sigma^2) g_{ss} \frac{d^2\mu_p}{d\rho^2} + 2 g_{pp} g_{ps} \frac{d\mu_p}{d\rho} P_{sp} + \mu_p g_{ps} g_{ss} \frac{d^2\mu_p}{d\rho^2} P_{sp} + (P_p g_{ps} + \sigma^2) g_{ss} \frac{d^2\mu_p}{d\rho^2} P_{sp}.
\end{align*} \tag{56}$$

with $\frac{d\mu_p}{d\rho}$ and $\frac{d^2\mu_p}{d\rho^2}$ given in (50), and $\frac{d\mu_p}{d\rho}$ and $\frac{d^2\mu_p}{d\rho^2}$ given in (19).

The proof for quasiconcavity of $R_{ss,n}$ is similar to that in Appendix A if we replace $\lambda_a$ in Appendix A by $\lambda_c$. The only difference is that we need to prove $(1 - \rho) \frac{d^2\lambda_c}{d\rho^2} - 2 \frac{d\lambda_c}{d\rho} > 0$ instead of the second equation of (54) in Appendix A. Base on (56), we have

$$(1 - \rho)\frac{d^2\lambda_c}{d\rho^2} - 2\frac{d\lambda_c}{d\rho} = \frac{2}{(1 - \rho)} \frac{d\lambda_c}{d\rho} P_{sp} \quad \text{and step (vi)} \tag{57}$$

in which step (vii) uses (56) and $\frac{d^2\mu_p}{d\rho^2} = \frac{2}{(1 - \rho)} \frac{d\mu_p}{d\rho}$ and step (viii) uses (55), $\frac{d\mu_p}{d\rho} > 0$ from (50), and $\frac{d^2\mu_p}{d\rho^2} > 0$ from (19).

**APPENDIX C**

**PROOF OF THEOREM 5**

$\xi_{w/o}$ given in (47) can be rewritten as

$$\xi_{w/o} = (1 - \rho)(\mu_p + 1)(P_{cgs} + P_{pgps} + \sigma^2) g_{ss} - (1 - \rho) [(\mu_p P_{cgs} + \sigma^2) - P_{pgps}] g_{ss} + (P_p g_{ps} + \sigma^2)(\mu_p + 1) g_{ss} \tag{57}$$

Defining $\chi = 2(\eta P_{gsh} + E_o - E_c)$ and $\psi = 2\eta P_{gsh}$, $P_c$ given in (1) can be represented as

$$P_c = \frac{1}{(1 - \rho)}(\chi - (1 - \rho)\psi) \tag{58}$$

Using (58) and applying some math manipulations, $\xi_{w/o}$ can be expressed as

$$\xi_{w/o} = \phi_a + \phi_b + \phi_c - \phi_d - \phi_e, \tag{59}$$

in which

$$\begin{align*}
\phi_a &= (\chi - (1 - \rho)\psi) \left(2^{\frac{2}{\lambda_a}} - P_{gsh} g_{ss} g_{sp}, \\
\phi_b &= (1 - \rho) \left(2^{\frac{2}{\lambda_a}} - P_{gsh} g_{ss} g_{sp}, \\
\phi_c &= (1 - \rho) \left(2^{\frac{2}{\lambda_a}} - P_{gsh} g_{ss} g_{sp}, \\
\phi_d &= (1 - \rho) \left(2^{\frac{2}{\lambda_a}} - P_{gsh} g_{ss} g_{sp}, \\
\phi_e &= (1 - \rho) \left(2^{\frac{2}{\lambda_a}} - P_{gsh} g_{ss} g_{sp}.
\end{align*} \tag{60}$$
The first order derivative of $\xi^{\omega/\alpha}$ is expressed by
\[
\frac{d^2 \xi^{\omega/\alpha}}{d\rho^2} = \frac{d\phi_a}{d\rho} + \frac{d\phi_b}{d\rho} - \frac{d\phi_c}{d\rho} - \frac{d\phi_d}{d\rho} - \frac{d\phi_e}{d\rho},
\]
where
\[
\begin{align*}
\frac{d\phi_a}{d\rho} &= (\psi(\rho t + 1) + (1 - \rho)P_1 \frac{d\rho t}{d\rho})g_{ss}g_{sp}, \\
\frac{d\phi_b}{d\rho} &= (-\psi(\rho t - 1)P_1 \frac{d\rho t}{d\rho} + (P_4 g_{sp} + \sigma^2)g_{sp}, \\
\frac{d\phi_c}{d\rho} &= (\psi(\rho t + 1)P_1 \frac{d\rho t}{d\rho} + (P_4 g_{sp} + \sigma^2)(\frac{(2ln2)t}{1-\rho} - 1))g_{ss}g_{sp}2^{\frac{2t}{\rho^2}}, \\
\frac{d\phi_d}{d\rho} &= (\psi(\rho t + 1)P_1 \frac{d\rho t}{d\rho} + (\psi+1)(\frac{(2ln2)t}{1-\rho} - 1))(P_4 g_{sp} + \sigma^2)g_{sp}2^{\frac{2t}{\rho^2}}.
\end{align*}
\]
 Define two functions $G(y)$ and $K(y)$ over $y \geq 0$ as $G(y) = \frac{(2ln2)y}{(1-\rho)\rho}$ and $K(y) = \frac{(2ln2)y}{(1-\rho)\rho} + \frac{(2ln2)y}{(1-\rho)\rho}$, respectively. We have
\[
G(\gamma T + t) = G(\gamma T) + G(t) + \frac{8(\ln 2)^2 \gamma T t}{(1-\rho)^3} \geq G(\gamma T) + G(t),
\]
(63)
\[
K(\gamma T + t) = K(\gamma T) + K(t) + \frac{8(\ln 2)^2 \gamma T t}{(1-\rho)^4} \geq K(\gamma T) + K(t).
\]
(64)
Then using $G(y)$ and $K(y)$, the second order derivatives of $(1-\rho)2^{\frac{2t}{\rho^2}}$ and $2^{\frac{2t}{\rho^2}}$ are written as $\frac{d^2}{d\rho^2}(1-\rho)2^{\frac{2t}{\rho^2}} = (\gamma T)t^2$ and $\frac{d^2}{d\rho^2}(2^{\frac{2t}{\rho^2}}) = K(\gamma T)t^2$. Using this, we have
\[
\begin{align*}
\frac{d^2 \phi_a}{d\rho^2} &= \left(\chi(\gamma T) - \psi(G(\gamma T))\right)g_{ss}g_{sp}2^{\frac{2t}{\rho^2}}, \\
\frac{d^2 \phi_b}{d\rho^2} &= G(\gamma T)(P_4 g_{sp} + \sigma^2)g_{sp}2^{\frac{2t}{\rho^2}}, \\
\frac{d^2 \phi_c}{d\rho^2} &= \left(\chi(\gamma T) - \psi(G(\gamma T))\right)2^{\frac{2t}{\rho^2} + \frac{8(\ln 2)^2 \gamma T t}{(1-\rho)^3}} - \left(\chi(\gamma t) - \psi(G(t))\right)(P_4 g_{sp} + \sigma^2 + 1)2^{\frac{2t}{\rho^2}}, \\
\frac{d^2 \phi_d}{d\rho^2} &= G(\gamma T)(P_4 g_{sp} + \sigma^2)g_{sp}2^{\frac{2t}{\rho^2} + \frac{8(\ln 2)^2 \gamma T t}{(1-\rho)^3}} - G(t)2^{\frac{2t}{\rho^2} + \frac{8(\ln 2)^2 \gamma T t}{(1-\rho)^3}}(P_4 g_{sp} + \sigma^2)2^{\frac{2t}{\rho^2}}, \\
\frac{d^2 \phi_e}{d\rho^2} &= G(\gamma T)(P_4 g_{sp} + \sigma^2)g_{sp}2^{\frac{2t}{\rho^2} + \frac{8(\ln 2)^2 \gamma T t}{(1-\rho)^3}} - G(t)2^{\frac{2t}{\rho^2} + \frac{8(\ln 2)^2 \gamma T t}{(1-\rho)^3}}(P_4 g_{sp} + \sigma^2)2^{\frac{2t}{\rho^2}}.
\end{align*}
\]
(65)
In the expressions for $\frac{d^2 \phi_a}{d\rho^2}$ and $\frac{d^2 \phi_b}{d\rho^2}$, both of them contain the item $\chi(K(y) - \psi(G(y))$, where $y$ takes the value $\gamma T$, $t$, or $(\gamma T + t)$. By checking $\chi(K(y) - \psi(G(y))$, it can be represented as the following
\[
\chi(K(y) - \psi(G(y)) = \chi(\frac{4(\ln 2)y}{(1-\rho)\rho} + \frac{(2\ln 2)y}{(1-\rho)^2}) - \psi(\frac{(2\ln 2)y}{(1-\rho)^2}),
\]
(66)
in which the last equality uses (58).

Then $\frac{d^2 \phi_a}{d\rho^2} - \frac{d^2 \phi_b}{d\rho^2}$ is expressed as
\[
\begin{align*}
\frac{d^2 \phi_a}{d\rho^2} - \frac{d^2 \phi_b}{d\rho^2} &= \left(\chi(K(\gamma T) - \psi(G(\gamma T))\right)g_{ss}g_{sp}2^{\frac{2t}{\rho^2} + \frac{8(\ln 2)^2 \gamma T t}{(1-\rho)^3}} \\
&\quad - \left(\chi(K(\gamma t) - \psi(G(t))\right)(P_4 g_{sp} + \sigma^2 + 1)2^{\frac{2t}{\rho^2}},
\end{align*}
\]
(67)
in which the first equality is from (65), and the second equality is from the first equality of (63), the first equality of (64), (10), (66), and (58).

Moreover, we have
\[
\frac{d^2 \phi_d}{d\rho^2} \geq \frac{d^2 \phi_e}{d\rho^2} = \left(\chi(\gamma t) - \psi(G(t))\right)(P_4 g_{sp} + \sigma^2 + 1)2^{\frac{2t}{\rho^2}},
\]
(68)
in which (ix) is from (65), and (xi) is from $G(\gamma T + t) \geq G(\gamma T) + G(t)$ and (10).

From (65) and $G(\gamma T + t) \geq G(\gamma T) + G(t)$, we have
\[
\begin{align*}
\frac{d^2 \phi_d}{d\rho^2} \geq \left(\chi(\gamma T) - \psi(G(t))\right)(P_4 g_{sp} + \sigma^2 + 1)2^{\frac{2t}{\rho^2}},
\end{align*}
\]
(69)
Combining (59), (67), (68), and (69), it can be seen that $\phi_a^{\omega/\alpha}$ is a strictly concave function.

REFERENCES


