New Viewpoint and Algorithms for Water-Filling Solutions in Wireless Communications

Chengwen Xing, Yindi Jing, Shuai Wang, Shaodan Ma, and H. Vincent Poor, *Fellow*, *IEEE*

***Abstract*—Water-filling solutions play an important role in the designs for wireless communications, e.g., transmit covariance matrix design. A traditional physical understanding is to use the analogy of pouring water over a pool with fluctuating bottom. Numerous variants of water-filling solutions have been discovered during the evolution of wireless networks. To obtain the solution values, iterative computations are required, even for simple cases with compact mathematical formulations. Thus, algorithm design is a key issue for the practical use of water-filling solutions, which however has been given marginal attention in the literature. Many existing algorithms are designed on a case-by-case basis for the variations of water-filling solutions and/or with overly complex logics. In this paper, a new viewpoint for water-filling solutions is proposed to understand the problem *dynamically* by considering changes in the increasing rates on different subchannels. This fresh viewpoint provides useful mechanism and fundamental information in finding the optimization solution values. Based on the new understanding, a novel and comprehensive method for practical water-filling algorithm design is proposed, which can be used for systems with various performance metrics and power constraints, even for systems with imperfect channel state information.**

I. Introduction

Water-filling solutions play a central role in the optimization of communication systems. They are undoubtfully among the most fundamental and important results in wireless communication designs, signal processing designs, and network optimizations including transceiver optimization, training optimization, resource allocation, and so on, e.g., [1]–[6]. Loosely speaking, optimal resource allocations for multi-dimensional communication systems usually lead to water-filling solutions. Over the past decade, wireless systems have evolved dramatically and exhibited a great variety of configurations with many different performance requirements and physical constraints, e.g., [7]. This diversity results in a rich body of variants of water-filling solutions [3]–[15], from single water-level ones to multiple water-level ones [4], from solutions for perfect channel state information (CSI) to robust ones such as cluster water-filling [13], and from constant water-level ones to cave-filling ones [15], [16].

C. Xing and S. Wang are with School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: chengwenxing@ieee.org and swang@bit.edu.cn).

Y. Jing is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 1H7, Canada (e-mail: yindi@ualberta.ca).

S. Ma is with the Department of Electrical and Computer Engineering, University of Macau, Macao SAR, China(e-mail: shaodanma@umac.mo).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA. (e-mail: poor@princeton.edu).

In existing work, the first step in obtaining a water-filling solution for an optimization problem is to find the Karush-Kuhn-Tucker (KKT) conditions and manipulate them into a recognizable format which is usually referred to as the water-filling solution. KKT conditions are necessary conditions for the optimization, and are also sufficient if the problem is convex [2]. While KKT conditions determine the optimal solutions, their initial formats are implicit and do not provide information in how to achieve the optimal solution values. Thus, sophisticated mathematical manipulations are needed to transform them into a water-filling structure.

As communication systems and optimization problems get more complicated, the corresponding KKT conditions also become more complicated, both in mathematical complexity and in the number of equations. Existing methods for this step become untenable. First, the large number of complicated KKT conditions hinder efficient manipulations and clear understanding of their physical meaning in terms of water-filling structure. Furthermore, the derived water-filling solutions may not have compact and systematic format to allow the development of effective algorithms.

Furthermore, the optimization design is not complete with the derived water-filling solutions as the solutions contain unknown parameters such as water levels. In other words, the solutions are still in *implicit form*. Thus, an important second step in obtaining the water-filling solution of an optimization problem is to find a practical algorithm. This step has not been given sufficient attention and in general has been ignored. Generally speaking, water-filling solutions consist of two major components, i.e., water level and water bottom, and a traditional imagery of puring is to pour water over a pool with different bottoms [1]. Based on this analogy, several practical water-filling algorithms have been proposed [16]–[21]. But they largely differ from each other and many have high complexity.

In this paper, we provide a new viewpoint on water-filling solutions. It has three major advantages: 1) it helps the understanding of water-filling results; 2) it avoids tedious and challenging manipulations of KKT conditions; and 3) it leads to efficient algorithms to find the solution values. Based on this new understanding, a unified method is proposed to find efficient water-filling algorithms for various complicated communication optimization problems. The main contributions are summarized as follows.

- We provide a novel understanding from a dynamic perspective for optimization problems with water-filling solutions. In contrast with the traditional approach, this viewpoint can avoid tedious manipulations of KKT conditions in deriving water-filling solutions and greatly simplify water-filling algorithm design.

- A standard and plausible notation used in water-filling solutions is the "+" operation where $x^+ \triangleq \max\{x, 0\}$. Its widely acknowledged physical meaning is that the resource (e.g., power) allocated to a subchannel must be nonnegative. However, this physical meaning, while intuitive, is not rigorous and can cause serious errors for complicated systems if misused as necessary condition or a guideline in mathematical derivations. Index based operations are introduced in the algorithm designs. The method deflects possible mistakes caused by the "+" operation, guarantees the optimality of the results, and simplifies the algorithm design.
- In addition to efficiency, the proposed method and the resulting algorithms are highly intuitive and understandable, and are also attractive from the implementation perspective. It is also compatible for extensions to complicated systems by using simple cases as building blocks.
- With the proposed method, we investigate a new class of communication optimizations with box constraints, where the allocated resource of each subchannel is bounded from both ends. Corresponding algorithms for the optimal solution values are proposed. Moreover, the algorithms can be extended to serve even more general problems and have a wide range of applications.
- Robust optimizations for wireless systems with CSI uncertainties are also studied. Algorithms for finding the solutions are again proposed for robust weighted mean-squared-error (MSE) minimization, robust capacity maximization, robust worst-MSE minimization, and robust minimum capacity maximization for multiple-input multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM) systems, the last two of which were largely open.

The remainder of this paper is organized as follows. In Section II, our new viewpoint on water-filling solutions is given, based on which an original algorithm is proposed to find the optimal solution. Following that, we investigate optimizations with box constraints in Section III. Then in Section IV, several extended optimizations are studied including problems with ascending sum-constraints, multiple water levels, and fairness considerations. Some numerical simulation results are given in Section V. Finally, the conclusions are drawn in Section VI.

II. A New Viewpoint of Water-filling Solutions

We consider a convex optimization problem of the following form:

$$\begin{aligned} \text{P1}: \max_{p_1,\cdots,p_K} \quad & \sum\nolimits_{k=1}^{K} f_k(p_k), \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \leq P, \quad p_k \geq 0, \end{aligned} \tag{1}$$

where $P > 0$ and the functions $f_k(\cdot)$ are real-valued, increasing, and strictly concave. Further assume that $f_k'(\cdot)$'s are continuous, where $f_k'(\cdot)$ denotes the first order derivative of $f_k(\cdot)$. Many optimization problems in wireless communications have this format or contain this problem as an essential part, for example, the power allocation problem in MIMO capacity maximization. It is known that the optimal solution of (1) has a water-filling structure. In what follows, we first explain the taditional treatment of this problem, then our new viewpoint and algorithm are elaborated along with the comparison of the two and examples.

A. Existing Treatment for Water-Filling Solutions

Traditionally, Lagrange multiplier method has been used for (1). The first step is to find the KKT conditions and from them to derive the water-filling solution of the problem in a compact format. As the objective function is a sum of decomposed concave functions and the constraints are linear, the problem is a convex one. Thus the KKT conditions are both necessary and sufficient. With straightforward calculations, the KKT conditions of (1) are

$$\begin{aligned} & f_k'(p_k) = \mu - \lambda_k, \\ & \mu \left(\sum\nolimits_{k=1}^{K} p_k - P \right) = 0, \\ & \lambda_k p_k = 0, \end{aligned} \tag{2}$$

where μ and λ_k are the Lagrange multipliers corresponding to the two constraint sets. By rewriting the KKT conditions, the solution has the following water-filling structure:

$$p_k = [g_k(\mu)]^+, \quad \text{and} \quad \sum\nolimits_{k=1}^{K} p_k = P, \tag{3}$$

where

$$g_k(\cdot) \triangleq \text{Inv}[f_k'](\cdot), \tag{4}$$

i.e., $g_k(\cdot)$ is the inverse function of $f_k'(\cdot)$. The Lagrange multiplier μ has the physical meaning of the water level. On the other hand, the function of λ_k's is implicit in this water-filling solution as they affect the solution through the "+" operation. We would like to highlight that the "+" operation results from rigorous mathematical derivations. While it can be explained intuitively by "the power level must be nonnegative," the "+" operation should not be added recklessly during the derivations merely due to this physical meaning. For more general problems, such practice can lead to mistake in the solution.

Another important step in using water-filling solutions in communications systems is to obtain the solution values, i.e., the values of p_k's of the solution in (3). It is a non-trivial step but has been somewhat neglected in the literature. In many cases, only brief discussion was provided and there is no implementable algorithm. The existing algorithms are usually for specific applications and a unified framework is missing.

To obtain the values of the p_k's from (3), a practical water-filling algorithm is needed. The major challenge is to find the index set of active subchannels with non-zero powers, i.e.,

$$\mathcal{S}_{active} = \{k | p_k > 0\}. \tag{5}$$

In general, all possible subchannel combinations need to be tried. For each of the $2^K - 1$ possibilities, the corresponding p_k-values can be found, and then the one with the highest objective function among the $2^K - 1$ possibilities is the optimal solution. But this method is obviously inefficient as the complexity is exponential in K. For simple settings with simple f_k-functions and fortunate parameter values, a natural ordering of the subchannels exists, and the algorithm can be re-designed to have lower complexity $\mathcal{O}(K \log K)$. But for most cases (e.g., Examples 1-5 in Section II-C), a natural ordering of the subchannels does not exist. A generic and efficient algorithm is still needed.

B. New Viewpoint and Algorithm

The traditional method for P1 as explained in the previous subsection has two major disadvantages. The first is the need of the transformation from KKT conditions to water-filling solutions. As the problem gets more general for more involved wireless systems and models, the transformation can become intractable. The second is the lack of general and effective algorithms in finding the values of the solution. In the following, from the perspective of a dynamic procedure, we give a new viewpoint on the solution of the optimization problem, which helps address both challenges.

Since $f_k(\cdot)$ is concave, $f_k'(\cdot)$ is a decreasing function meaning that the increasing rate of $f_k(p_k)$ decreases as p_k increases. The optimization problem P1 aims at allocating the total power P over a series of functions, i.e., $f_k(p_k)$'s. We can see this problem as dividing the available power P into a large number of small portions and the power is allocated portion by portion. For each portion, we should choose the subchannel whose $f_k(\cdot)$ has the maximum increasing rate to maximize the total of $f_k(\cdot)$'s. As the increasing rate of this f_k-function decreases when a resource portion is added to it, after getting a certain amount of power portions, its increasing rate may become smaller than another subchannel. In this case, a new subchannel will have the fastest increasing rate and the next power portion should be added to this new subchannel. This procedure repeats until all resource portions have been allocated. When the resource allocation stops, the functions that are allocated with nonzero powers will have the same increasing rate. Some subchannels may never get any power portion if their increasing rates are never the highest.

The result discussed above is presented in the following claim with rigorous proof.

Lemma 1 *The following conditions are both necessary and sufficient for the optimal solution of P1:*

$$\begin{cases} f_k'(p_k) = f_j'(p_j) & \text{for } k, j \in \mathcal{S}_{active}; \\ f_j'(0) \leq f_k'(p_k) & \text{for } k \in \mathcal{S}_{active} \text{ and } j \notin \mathcal{S}_{active}; \\ \sum_{k=1}^{K} p_k = P. \end{cases} \quad (6)$$

Proof: We first prove the necessity part by contradiction. The necessity of the last line of (6) is obvious and has been proved in many existing work. Thus the proof is omitted here. Denote the optimal solution of P1 as $p_1^*, \cdots, p_K^*$. Assume without loss of generality that $p_1^*, p_2^* > 0$ (i.e., $1, 2 \in \mathcal{S}_{active}$) but $f_1'(p_1^*) > f_2'(p_2^*)$. Since f_1' and f_2' are continuous, there exists an δ with $0 < \delta < p_2^*$ such that $f_1'(p_1^* + x) > f_2'(p_2^* - x)$ for $0 < x \leq \delta$. Thus

$$\begin{aligned} & f_1(p_1^* + \delta) + f_2(p_2^* - \delta) \\ &= f_1(p_1^*) + f_2(p_2^*) + \int_0^\delta \left[f_1'(p_1^* + x) - f_2'(p_2^* - x)\right] dx \\ &> f_1(p_1^*) + f_2(p_2^*). \end{aligned}$$

This shows that the new solution $\{p_1^* + \delta, p_2^* - \delta, p_3^*, \cdots, p_K^*\}$ (which satisfies all constraints by construction) is better than $p_1^*, p_2^*, p_3^*, \cdots, p_K^*$, which contradicts the assumption. This proves that the first line of (6) is necessary.

Similarly, to prove that the second line of (6) is necessary, assume without loss of generality that $p_1^* > 0, p_2^* = 0$ (i.e., $1 \in \mathcal{S}_{active}$ and $2 \notin \mathcal{S}_{active}$) but $f_2'(0) > f_1'(p_1^*)$. Since f_2 is strictly concave and f_2' is continuous, there exists an δ with $0 < \delta < p_1^*$ such that $f_1'(p_1^* - x) < f_2'(x)$ for $0 < x \leq \delta$. Thus

$$\begin{aligned} & f_1(p_1^* - \delta) + f_2(\delta) \\ &= f_1(p_1^*) + f_2(0) + \int_0^\delta \left[f_2'(x) - f_1'(p_1^* - x)\right] dx \\ &> f_1(p_1^*) + f_2(0). \end{aligned}$$

This says that the solution $\{p_1^* - \delta, \delta, p_3^*, \cdots, p_K^*\}$ is better and thus leads to a contradiction.

For the sufficiency, it is enough to show that a solution satisfying (6) is a local maximum. Since P1 is a convex optimization, its local maximum is unique and is the global maximum. Let $\{p_1^*, \cdots, p_K^*\}$ be the solution satisfying (6) and consider a solution $\{p_1, \cdots, p_K\}$ in the vicinity of it. Define $\mathcal{S}_+ \triangleq \{k | p_k > p_k^*\}$ and $\mathcal{S}_- \triangleq \{k | p_k < p_k^*\}$. Notice that $\mathcal{S}_- \cap \mathcal{S}_{inactive}^* = \emptyset$, where $\mathcal{S}_{inactive}^* = \{k | p_k^* = 0\}$. Thus

$$\begin{aligned} \sum_{k=1}^{K} f_k(p_k) = \sum_{k=1}^{K} f_k(p_k^*) + \sum_{k \in \mathcal{S}_+} \int_0^{p_k - p_k^*} f_k'(p_k^* + x_k) dx_k \\ - \sum_{\hat{k} \in \mathcal{S}_-} \int_0^{p_{\hat{k}}^* - p_{\hat{k}}} f_{\hat{k}}'(p_{\hat{k}}^* - \hat{x}_{\hat{k}}) d\hat{x}_{\hat{k}}. \end{aligned}$$

From the conditions on f_k's and the assumption that $\{p_1^*, \cdots, p_K^*\}$ satisfies (6), we have

$$f_k'(p_k^* + x) < f_k'(p_k^*) \leq f_{\hat{k}}'(p_{\hat{k}}^*) < f_{\hat{k}}'(p_{\hat{k}}^* - \hat{x}), \quad (7)$$

for all $k \in \mathcal{S}_+$, $\hat{k} \in \mathcal{S}_-$, $x_k \in (0, p_k - p_k^*)$, $\hat{x}_{\hat{k}} \in (0, p_{\hat{k}}^* - p_{\hat{k}})$. Also, since $\sum_{k=1}^{K} p_k^* = P \geq \sum_{k=1}^{K} p_k$, we have

$$\sum_{k \in \mathcal{S}_+} (p_k - p_k^*) \leq \sum_{\hat{k} \in \mathcal{S}_-} (p_{\hat{k}}^* - p_{\hat{k}}). \quad (8)$$

By combining (7) and (8), it can be concluded that $\sum_{k=1}^{K} f_k(p_k) < \sum_{k=1}^{K} f_k(p_k^*)$, and thus $\{p_1^*, \cdots, p_K^*\}$ is a local maximum.[1] ■

From (6), we see that the value of $f_k'(p_k)$ for $k \in \mathcal{S}_{active}$, denoted as μ, is the increasing rate for the optimal power allocation result. The allocated power on the subchannels can also be represented as functions of μ:

$$\begin{cases} p_k = g_k(\mu) & \text{for } k \in \mathcal{S}_{active} \\ p_k = 0 & \text{for } k \notin \mathcal{S}_{active} \end{cases}, \quad (9)$$

[1]The lemma can also be proved by showing that (6) is equivalent to the KKT conditions, which are necessary and sufficient for P1. But here we use a direct proof to help illustrate the proposed new viewpoint and avoid unnecessary dependence on existing water-filling results.

Algorithm 1 Proposed water-filling algorithm for P1.

1: $\mathcal{I}_k = 1$ for $k = 1, \cdots, K$;
2: Calculate μ and p_k's using Eqn. (11);
3: **while** length(find($\{p_k\} < 0$)) > 0 **do**
4:     Find $\mathcal{S}_{inactive} = \{k | p_k \leq 0\}$;
5:     Set $\mathcal{I}_k = 0$ for $k \in \mathcal{S}_{inactive}$;
6:     Calculate μ and p_k's using Eqn. (11);
7: **end while**
8: **return** p_k's

where $g_k(\cdot)$ is defined in (4). From the total power constraint,

$$P = \sum\nolimits_{k \in \mathcal{S}_{active}} g_k(\mu), \tag{10}$$

based on which μ can be solved when the set of active subchannels $\mathcal{S}_{active}$ is known.

As explained in the previous subsection. The main difficulty of finding the solution values is to find $\mathcal{S}_{active}$. We propose the use of index operations $\mathcal{I}_k$'s to conquer this difficulty. When subchannel k is allocated nonzero power, $\mathcal{I}_k = 1$, otherwise $\mathcal{I}_k = 0$. Based on these indices, (9) and (10) are rewritten as

$$p_k = g_k(\mu)\mathcal{I}_k \text{ and } P = \sum\nolimits_{k=1}^{K} g_k(\mu)\mathcal{I}_k. \tag{11}$$

With this result, we present a new water-filling algorithm for P1 in **Algorithm 1**.

In the first step of **Algorithm 1**, all subchannels are initialized as active and in the second step, the corresponding increasing rate and subchannel powers are calculated. As the computations of p_k's do not consider the constraints that $p_k \geq 0$, it may appear that $p_k < 0$ for some k. In this case, the corresponding index $\mathcal{I}_k$ will be set to zero and this subchannel will be allocated zero-power in the next round. The procedure continues until all active subchannels are allocated nonnegative powers.

Lemma 2 *Algorithm 1 converges and achieves the optimal solution of P1.*

Proof: Since for each iteration in Algorithm 1, the new set for $\mathcal{S}_{active} = \{k | p_k > 0\}$ is either the same as the previous $\mathcal{S}_{active}$ (thus the algorithm terminates) or shrinks to a subset of the previous $\mathcal{S}_{active}$. As the size of the initial set is K, it is obvious that the algorithm converges within K iterations.

Now we prove that Algorithm 1 converges to the optimal solution of P1. First, since $P > 0$, at any iteration, it is impossible to have $p_k \leq 0$ for all k. In other words, there exists a k such that $p_k > 0$. Let $\{p_1, \cdots, p_K\}$ be the solution found by Algorithm 1 at the mth iteration. From Step 2 and Step 6, it is obvious that the solution satisfies the first and last conditions of (6). For any $j \notin \mathcal{S}_{active}$, we have $p_j < 0$ in one of the previous iterations. Denote the iteration round for $p_j < 0$ as m'. Thus from (11), $p_j = g_j(\mu^{(m')}) < 0$, from which $\mu^{(m')} > f_j'(0)$, where $\mu^{(m')}$ is the achieved increasing rate at the m'th iteration. Notice that $\mu^{(m')} = f_k'(p_k^{(m')})$ for subchannel k in the active set of the m'th iteration. With the proposed algorithm, subchannel j is removed by setting $p_j = 0$, and in the next iteration, the sum power available for the remaining active subchannels decreases. The achieved increasing rate for this new iteration is higher, i.e., $\mu^{(m')} < \mu^{(m'+1)}$. Denote the overall iteration number for the algorithm as m. Since $m' \leq m$, we have $f_j'(0) < \mu^{(m')} \leq \mu^{(m)} = f_k'(p_k)$ for $k \in \mathcal{S}_{active}$. This proves that the solution found by the algorithm also satisfies the second condition of (6). As (6) is proved to be sufficient for the optimal solution in Lemma 1, the solution found by Algorithm 1 is thus the optimal one. ■

C. Comparison and Application Examples

The proposed new method, including the viewpoint and the algorithm, does not require manipulation of the KKT conditions into a format of water-filling solutions. Further, the proposed algorithm is general and has low-complexity with the worst-case number of iterations being $K-1$. On average, the number of iterations can be much smaller than $K-1$ since the proposed algorithm allows multiple channels to be made inactive in each iteration as long as their positivity constraints cannot be satisfied. For the traditional scheme, in general, all possible subsets of active subchannels need to be tested, whose complexity is exponential in K. For special cases when an ordering among the subchannel exists, the complexity can be reduced to $\mathcal{O}(K\log(K))$, which is still higher than the complexity of the proposed one. In what follows, examples are provided to better elaborate the difference and advantages of the proposed method.

Example 1: A general weighted sum capacity maximization problem has the following form:

$$\begin{aligned} &\max_{p_1, \cdots, p_K} \quad \sum\nolimits_{k=1}^{K} w_k \log(b_k + a_k p_k) \\ &\text{s.t.} \quad \sum\nolimits_{k=1}^{K} p_k \leq P, \quad p_k \geq 0, \end{aligned} \tag{12}$$

where w_k, b_k and a_k are arbitrary non-negative parameters.

With our proposed scheme, we first obtain from the objective function in (12)

$$g_k(u) = \frac{w_k}{\mu} - \frac{b_k}{a_k}. \tag{13}$$

Then the solution values can be found by **Algorithm 1** within $K-1$ iterations. Specifically, from (11),

$$\mu = \frac{\sum_{\mathcal{I}_k=1} w_k}{P + \sum_{\mathcal{I}_k=1} \frac{b_k}{a_k}}. \tag{14}$$

The calculations in Step 2 and Step 6 can be achieved straightforwardly using (13) and (14).

With the traditional scheme, via calculations, the following water-filling solution is obtained:

$$p_k = \left(\frac{w_k}{\mu} - \frac{b_k}{a_k}\right)^+, \quad \sum\nolimits_{k=1}^{K} p_k = P. \tag{15}$$

Though in compact neat form, to find the values of the optimal p_k's is not self-explanatory. All possible active subchannel sets need to be tried to find the best one. In [18], efficient algorithms were proposed for two special cases: 1) $b_k = 1, w_k = 1$ and 2) $b_k = 1$, both w_i and $a_i w_i$ are in decreasing order. For these two cases, an order of the subchannels for their priority of getting non-zero power exists. But for the general case, an order does not exist (for example, for some weight values, a_i's and $(w_i a_i)$'s cannot be ordered decreasingly at the same time) and the algorithms in [18] cannot be applied.

Example 2: A general weighted MSE minimization problem can be written in the following form:

$$\begin{aligned}\min_{p_1,\cdots,p_K} \quad & \sum\nolimits_{k=1}^{K} \frac{w_k}{b_k + a_k p_k} \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \le P, \quad p_k \ge 0,\end{aligned} \tag{16}$$

where w_k, b_k and a_k are arbitrary non-negative parameters.

With the proposed scheme, we first obtain from the problem

$$g_k(u) = \sqrt{\frac{w_k}{a_k \mu}} - \frac{b_k}{a_k}. \tag{17}$$

Similarly, **Algorithm 1** can be used to find the solution values. Specifically, from (11),

$$\mu = \frac{\sum_{\mathcal{I}_k=1} \sqrt{\frac{w_k}{a_k}}}{P + \sum_{\mathcal{I}_k=1} \frac{b_k}{a_k}}. \tag{18}$$

(17) and (18) can be used straightforwardly for the calculations in Steps 2 and 6.

With the traditional scheme, via calculations, the following water-filling solution is obtained as the first step:

$$p_k = \left(\sqrt{\frac{w_k}{a_k \mu}} - \frac{b_k}{a_k}\right)^+, \quad \sum\nolimits_{k=1}^{K} p_k = P. \tag{19}$$

The same difficulty as in Example 1 appears here. Though (19) is in compact neat form, it is unclear how to find the values of the optimal solution from it. In general all possible active subchannel sets need to be tried to find the best one whose complexity is exponential in K. Ordering of the subchannels is only possible with stringent ordering conditions on the parameters, e.g., $\sqrt{w_k/a_k}$ and a_k/b_k can be ordered decreasingly simultaneously.

Example 3: The capacity maximization for dual-hop MIMO amplify-and-forward relaying networks can be casted as follows [21], [24]:

$$\begin{aligned}\min_{p_1,\cdots,p_K} \quad & \sum\nolimits_{k=1}^{K} w_k \log\left(1 - \frac{a_k b_k p_k}{1 + b_k p_k}\right) \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \le P, \quad p_k \ge 0,\end{aligned} \tag{20}$$

where w_k, b_k are nonnegative and $0 < a_k < 1$.

With our proposed scheme, we can obtain from the objective function of the problem

$$g_k(\mu) = \frac{\sqrt{a_k^2 + 4w_k(1-a_k)a_k b_k/\mu} - (2 - a_k)}{2(1-a_k)b_k}.$$

Then the solution values can be found by **Algorithm 1**. But for this case, to find the value of μ (for Steps 2 and 6), numerical bisection search is needed to solve the following equation

$$\sum_{\mathcal{I}_k=1} \frac{\sqrt{a_k^2 + 4w_k(1-a_k)a_k b_k/\mu} - (2 - a_k)}{2(1-a_k)b_k} = P.$$

With the traditional scheme, via some calculations, the following water-filling solution is obtained as the first step:

$$\begin{aligned}& p_k = \left(\frac{a_k - 2 + \sqrt{a_k^2 + 4w_k(1-a_k)a_k b_k/\mu}}{2(1-a_k)b_k}\right)^+. \\ & \sum\nolimits_{k=1}^{K} p_k = P.\end{aligned}$$

But algorithms to find the water-filling solution values were not explicitly provided in existing literature.

Example 4: A weighted mutual information maximization problem for the training design can be written in the following format [22]:

$$\begin{aligned}\max_{p_1,\cdots,p_K} \quad & \sum\nolimits_{j=1}^{J} \sum\nolimits_{k=1}^{K} w_{k,j} \log(a_k c_j + b_k d_j p_k), \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \le P, \quad p_k \ge 0.\end{aligned} \tag{21}$$

To use the proposed scheme, we first get from the objective function

$$f_k'(p_k) = \sum\nolimits_{j=1}^{J} \frac{w_{k,j} b_k d_j}{a_k c_j + b_k d_j p_k}. \tag{22}$$

Due to the complexity of f_k', the function $g_k(\mu)$ does not have an explicit closed-form. But **Algorithm 1** can still be used to find the solution values by calculating μ and p_k's numerically in Steps 2 and 6 using (4), (11), and (22).

With the traditional scheme, via calculations, the following water-filling solution is obtained in the first step:

$$\begin{aligned}& \sum\nolimits_{j=1}^{J} \frac{w_{k,j} b_k d_j}{a_k c_j + b_k d_j p_k} = \mu - \lambda_k, \\ & \sum\nolimits_{k=1}^{K} p_k \le P, \quad \lambda_k p_k = 0.\end{aligned}$$

Unlike Examples 1-3, the KKT conditions for this example cannot be written in a compact water-filling solution form by using the "+" operation and no efficient algorithm was available in the literature to find the solution values.

Example 5: A weighted MSE minimization problem for training optimization can be formulated as follows:

$$\begin{aligned}\min_{p_1,\cdots,p_K} \quad & \sum\nolimits_{j=1}^{J} \sum\nolimits_{k=1}^{K} \frac{w_{k,j}}{a_k c_j + b_k d_j p_k} \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \le P, \quad p_k \ge 0,\end{aligned} \tag{23}$$

To use the proposed scheme, we first get from the objective function

$$f_k'(p_k) = \sum\nolimits_{j=1}^{J} \frac{w_{k,j} b_k d_j}{(a_k c_j + b_k d_j p_k)^2}.$$

Again, the function $g_k(\mu)$ does not have an explicit closed-form, but **Algorithm 1** can still be used to find the solution values by calculating μ and p_k's numerically in Steps 2 and 6.

With the traditional scheme, similar to Example 4, the KKT conditions can be obtained but a compact water-filling solution form has not been found with the "+" operation, nor have efficient algorithms been proposed to find the solution values in the literature.

Algorithm 2 Proposed algorithm under arbitrary lower-bound constraints.

```
1: I_k = 1 for k = 1, ..., K;
2: Calculate μ and p_k's using Eqn. (26);
3: while length(find({p_k} < {γ_k})) > 0 do
4:     Find S_inactive = {k | p_k ≤ γ_k};
5:     Set I_k = 0 and p_k = γ_k for k ∈ S_inactive;
6:     Calculate μ and p_k's using Eqn. (26);
7: end while
8: return p_k's.
```

D. Problems with Arbitrary Lower-Bound Constraints

In this subsection, we consider the extension of the optimization problem P1 with arbitrary lower bounds on the subchannel powers:

$$\begin{aligned} \mathrm{P1.1}: \max_{p_1,\cdots,p_K} \quad & \sum\nolimits_{k=1}^{K} f_k(p_k) \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \le P, \quad p_k \ge \gamma_k, \quad k = 1,\cdots,K, \end{aligned} \tag{24}$$

where $P > 0$ and $f_k(\cdot)$'s are real-valued, increasing, and strictly concave functions with continuous derivatives.

In P1.1, each subchannel is limited with a non-negative lower bound for its power, while for P1, the lower bounds are zero for all subchannels. For this more general case, define the active set $\mathcal{S}_{active}$ as the set of subchannels whose powers are higher than their lower bounds, i.e.,

$$\mathcal{S}_{active} \triangleq \{k | p_k > \gamma_k\}.$$

The following lemma is obtained.

Lemma 3 *The following conditions are both necessary and sufficient for the optimal solution of P1.1:*

$$\begin{cases} f_k'(p_k) = f_j'(p_j) \, for \; k, j \in \mathcal{S}_{active}; \\ f_j'(\gamma_j) \le f_k'(p_k) \, for \; k \in \mathcal{S}_{active} \; and, j \notin \mathcal{S}_{active}; \\ \sum_{k=1}^{K} p_k = P. \end{cases} \tag{25}$$

Proof: The proof is very similar to that of Lemma 1, thus omitted. ■

For the algorithm design, the index operation $\mathcal{I}_k$ is introduced as follows: $\mathcal{I}_k = 1$ when the power of subchannel k is larger than its lower bound, i.e., $p_k > \gamma_k$; otherwise $\mathcal{I}_k = 0$. Let $\mu = f_k'(p_k)$ for $k \in \mathcal{S}_{active}$, which is the increasing rate for active subchannels. Via similar studies to those in Section II-B, the optimal solution of P1.1 can be represented as follows:

$$\begin{cases} p_k = g_k(\mu)\mathcal{I}_k + \gamma_k(1 - \mathcal{I}_k) \\ P = \sum_{k=1}^{K} \left[g_k(\mu)\mathcal{I}_k + \gamma_k(1 - \mathcal{I}_k) \right] \end{cases}. \tag{26}$$

Notice that (11) is a special case of (26) where $\gamma_k = 0$. **Algorithm 2** is proposed to find the solution values for P1.1.

Lemma 4 ***Algorithm 2*** *converges and achieves the optimal solution of P1.1.*

Proof: The proof is similar to that of Lemma 2, thus omitted.■

In each iteration of **Algorithm 2**, subchannels whose powers are less than their required lower bounds are removed from the iteration (i.e., are put in the inactive set) and their powers are enforced to be the corresponding lower bounds, i.e., $p_k = \gamma_k$. Since these subchannels are allocated smaller powers than their lower bounds, their increasing rates at the lower bounds γ_k are smaller than other subchannels. After being removed, fewer power resources are available for the remaining active subchannels. After power allocation among the remaining subchannels in Step 6, the powers of the active subchannels decrease, and thus their increasing rates will increase. Therefore, the removed subchannels cannot enter the competition for power in future iterations. This explains the convergence and optimality of the algorithm intuitively.

E. Discussions on More General Cases

The new viewpoint and method can be extended to solve more general optimization problems in wireless communications. Consider the following convex optimization problem:

$$\begin{aligned} \mathrm{P2}: \max_{p_1,\cdots,p_K} \quad & \sum\nolimits_{k=1}^{K} f_k(p_k) \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \le P, \\ \text{and} \quad & h_l(p_k) \le 0, \quad k = 1,\cdots,K, \; l = 1,\cdots,L, \end{aligned} \tag{27}$$

where $P > 0$ and $f_k(\cdot)$'s are real-valued, increasing, and strictly concave functions with continuous derivatives.

The difference of P2 to the original one P1 is in the constraints $h_l(p_k)$'s. As P2 is convex, the following KKT conditions are necessary and sufficient for the optimal solution [2]:

$$\begin{aligned} & f_k'(p_k) = \mu + \sum\nolimits_l \lambda_l h_l'(p_k), \quad \mu \left(\sum\nolimits_{k=1}^{K} p_k - P \right) = 0, \\ & \lambda_l h_l(p_k) = 0, \quad \mu \ge 0, \quad \lambda_k \ge 0, \end{aligned} \tag{28}$$

where μ and λ_l's are the Lagrange multipliers corresponding to the sum power constraint and per-subchannel constraints, respectively.

There are no unified methods or efficient algorithms to find the solution values of P2 in the literature. By following the ideas proposed in previous subsections, we can solve this challenging problem by considering two situations: 1) all conditions $h_l(p_k)$'s are inactive (i.e., not satisfied with equality) and 2) at least one of $h_l(p_k)$'s is active (i.e., satisfied with equality). The first situation leads to the same solution as P1. For the second one, the results for P1 can be applied for the power allocation among subchannels with inactive $h_l(p_k)$'s and solutions for subchannels with active $h_l(p_k)$'s can be found by solving $h_l(p_k) = 0$. In the following sections, we will solve the generalized problem considering several different cases.

III. Problem with Box Constraints

In this section, we consider a special case of P2 in which $L = 1$ and $h_l(p_k) = (p_k - \gamma_k)(p_k - \tau_k)$. Equivalently, the optimization problem is as follows:

$$\mathrm{P3}: \max_{p_1,\cdots,p_K} \quad \sum\nolimits_{k=1}^{K} f_k(p_k)$$

$$\text{s.t.} \quad \sum_{k=1}^{K} p_k \le P,$$
$$\text{and} \quad \gamma_k \le p_k \le \tau_k, \quad k = 1, \cdots, K, \tag{29}$$

where $P > 0$, $\gamma_k \le \tau_k$, and $f_k(\cdot)$'s are real-valued, increasing, and strictly concave functions with continuous derivatives.

A. Two Algorithms Built on Finding Subchannel Sets

Similar to the previous section, we can see this problem as dividing the available power P into infinitesimally small portions δ_p and allocating them portion by portion. At the start of the allocation, Subchannel k must have γ_k to satisfy the lower bound constraint. For each remaining portion, we should choose the subchannel whose $f_k(\cdot)$ has the maximum increasing rate i.e., $f'_k(p_k)$, and whose power has not reached its upper bound to maximize the total of $f_k(\cdot)$'s. As the increasing rate of f_k decreases when a power portion is added to it, after adding a portion to the subchannel with the maximum increasing rate, e.g., Subchannel k, its increasing rate may become smaller to another subchannel. In this case, a new subchannel j with the fastest increasing rate will have the next power portion. Otherwise, Subchannel k gets the next power portion if it still has the maximum $f'_k(p_k + \delta_p)$. This procedure repeats until all power portions have been allocated. Some subchannels may never get any extra power portion than the original lower bounds when their increasing rates are never the highest. Some subchannels may have the highest increasing rates but cannot get more power due to their upper bound constraints. When the allocation stops, subchannels which do not have active bounds must have the same increasing rate.

For a given feasible solution $\{p_1, \cdots, p_K\}$, denote

$$\mathcal{S}_l \triangleq \{k | p_k = \gamma_k\}, \quad \mathcal{S}_u \triangleq \{k | p_k = \tau_k\},$$
$$\mathcal{S}_{active} \triangleq \{k | \gamma_k < p_k < \tau_k\}, \tag{30}$$

which are the index sets of subchannels whose power values equal their lower bounds, upper bounds, and in-between the two bounds (i.e., active subchannels), respectively. They are also sets of subchannels with active lower bounds, active upper bounds, and no active bounds. The following lemma provides the sufficient and necessary condition on the optimal solution of P3.

Lemma 5 *The following conditions are both necessary and sufficient for the optimal solution of P3:*

$$\begin{cases} f'_k(p_k) = f'_j(p_j) \text{ for } k, j \in \mathcal{S}_{active}; \\ f'_j(\gamma_j) \le f'_k(p_k) \text{ for } k \in \mathcal{S}_{active} \text{ and } j \in \mathcal{S}_l; \\ f'_j(\tau_j) \ge f'_k(p_k) \text{ for } k \in \mathcal{S}_{active} \text{ and } j \in \mathcal{S}_u; \\ \sum_{k=1}^{K} p_k = P. \end{cases} \tag{31}$$

Proof The proof is similar to that of Lemma 1 with the following two changes: 1) the lower bounds change from 0 to γ_k's and 2) new upper bounds are added. Details are omitted to save space. ∎

The physical meaning of (31) is as follows. At the optimal solution, subchannels with inactive bounds have the same the increasing rate $f'_k(p_k)$, which is also denoted as μ. Subchannels with active lower bounds have lower increasing rates than μ and subchannels with active upper bounds have higher increasing rates than μ.

Algorithm 3 The first proposed algorithm for P3.

```
Perform Algorithm 2;
while length(find({p_k} > {τ_k})) > 0 do
    Find S_u = {p_k ≥ τ_k} and set p_k = τ_k for k ∈ S_u;
    Find S_other = {p_k < τ_k};
    Update P ← P − Σ_{k∈S_u} τ_k;
    Perform Algorithm 2 for subchannels in S_other with
    the updated total power P;
end while
return p_k's.
```

Algorithm 4 The second proposed balanced algorithm for P3.

```
I_k = 1 and J_k = 1 for k = 1, ···, K.
Calculate μ and p_k's using Eqn. (32).
while length(find(p_k < γ_k))+length(find(p_k > τ_k))>0 do
    Find S_l = {k|p_k ≤ γ_k} and set p_k = γ_k, I_k = 0 for
    k ∈ S_l;
    Find S_u = {k|p_k ≥ τ_k} and set p_k = τ_k, J_k = 0 for
    k ∈ S_u;
    Calculate μ and p_k's using Eqn. (32);
end while
return p_k's
```

Based on the viewpoint and conditions for the optimal solution of P3, we propose **Algorithm 3** to find the solution values by using Algorithm 2 as a building block. The idea is to first consider the lower bound constraints only and use Algorithm 2 to find the corresponding solution. Then the subchannels whose power values are larger or the same as their upper bounds are re-set as their upper bounds, and are removed from the next iteration. In the next iteration, power is allocated among the remaining subchannels using Algorithm 2 again. The process continues until the powers of all subchanels are smaller than their upper bounds at an iteration.

Algorithm 3 does not have balanced treatment between the lower bound constraints and the upper bound constraints. While subchannels that reach or violate their upper bound constraints are removed during the iterations, the ones reaching or violating their lower constraints stay in the 'while' loop and participate in the power allocation procedure with Algorithm 2. Another algorithm symmetrical to Algorithm 3 can also be designed by switching the roles of the lower and upper bound constraints.

Next, we consider both constraints jointly and symmetrically. Based on the aforementioned discussions, the key task is to determine the sets $\mathcal{S}_u$, $\mathcal{S}_l$, and $\mathcal{S}_{active}$ defined in (30). We introduce two sets of indices $\mathcal{I}_k$'s and $\mathcal{J}_k$'s as follows:

$$\mathcal{I}_k = \begin{cases} 1 & \text{if } p_k > \gamma_k \\ 0 & \text{otherwise} \end{cases}, \quad \mathcal{J}_k = \begin{cases} 1 & \text{if } p_k < \tau_k \\ 0 & \text{otherwise} \end{cases},$$

where p_k is the power allocated to subchannel k, $\mathcal{I}_k$ indicates whether the power of subchannel k is larger than its lower bound constraint and $\mathcal{J}_k$ indicates whether it is smaller than its upper bound constraint. For the index tuple $(\mathcal{I}_k, \mathcal{J}_k)$, $(1, 1)$ means the subchannel is an active one and neither constraints

Algorithm 5 The third proposed bisection search based algorithm for P3.

Initialize σ, $\mu_{\min}$ and $\mu_{\max}$.
Let $\mu = (\mu_{\min} + \mu_{\max})/2$. Compute p_k's using the first formula of Eqn. (32).
Find $\mathcal{S}_u = \{p_k > \tau_k\}$ and set $p_k = \tau_k$ for $k \in \mathcal{S}_u$;
Find $\mathcal{S}_l = \{p_k < \tau_k\}$ and set $p_k = \gamma_k$ for $k \in \mathcal{S}_l$;
while $|\sum_k p_k - P| > \sigma$ **do**
  if $\sum_k p_k > P$ **then**
    $\mu_{\max} = (\mu_{\min} + \mu_{\max})/2$
  else
    $\mu_{\min} = (\mu_{\min} + \mu_{\max})/2$
  end if
  Let $\mu = (\mu_{\min} + \mu_{\max})/2$ and compute p_k's using the first formula of Eqn. (32).
  Find $\mathcal{S}_u = \{p_k > \tau_k\}$ and set $p_k = \tau_k$ for $k \in \mathcal{S}_u$;
  Find $\mathcal{S}_l = \{p_k < \tau_k\}$ and set $p_k = \gamma_k$ for $k \in \mathcal{S}_l$;
end while
return p_k's.

is tight; $(1, 0)$ means the subchannel belongs to $\mathcal{S}_u$; and $(0, 1)$ means the subchannel belongs to $\mathcal{S}_l$. Similar to the previous section, let μ be the increasing rate of the active subchannels, and we have the following necessary conditions for P3 from (31):

$$\begin{cases} p_k = g_k(\mu)\mathcal{I}_k\mathcal{J}_k + \gamma_k(1 - \mathcal{I}_k) + \tau_k(1 - \mathcal{J}_k), \\ P = \sum_{k=1}^{K} \left[g_k(\mu)\mathcal{I}_k\mathcal{J}_k + \gamma_k(1 - \mathcal{I}_k) + \tau_k(1 - \mathcal{J}_k)\right] \end{cases} \tag{32}$$

Algorithm 4 is proposed which follows the idea in Algorithm 2 with extensions for both lower and upper bound constraints.

By using results in Lemma 5 and following the proof in Lemma 4, the convergence and optimality of Algorithms 3 and 4 can be proved.

Lemma 6 ***Algorithm 3*** *and* ***Algorithm 4*** *converge and achieve the optimal solution of P3.*

Proof The detailed proof is similar to that of Lemma 4 and is thus omitted to save space. ■

B. Two Algorithms Built on Finding the Final Increasing Rate

In this subsection, two new algorithms are proposed, which are constructed by finding the final increasing rate μ. We first propose a complex but general one in **Algorithm 5**, where bisection search is used. It is easy to understand and implement, but suffers high complexity and numerical accuracy limitations.

The next algorithm, **Algorithm 6**, uses a more efficient method to find the final increasing rate μ. Firstly, the subchannels are ordered decreasingly based on their increasing rates at the power upper bounds $f'_k(\tau_k)$'s such that

$$f'_{\sigma_1}(\tau_{\sigma_1}) \geq \cdots \geq f'_{\sigma_K}(\tau_{\sigma_K}). \tag{33}$$

From the results in Lemma 5, it can be shown that the subchannel with a higher $f'_k(\tau_k)$ has higher priority to achieve its upper bound. In other words, if at the optimal solution $p_{\sigma_i} = \tau_{\sigma_i}$, then $p_{\sigma_j} = \tau_{\sigma_j}$ for all $j < i$. Thus, in finding the optimal solution, we can consider the cases of $p_{\sigma_{[1:i]}} = \tau_{\sigma_{[1:i]}}$ and $p_{\sigma_{[i+1:K]}} < \tau_{\sigma_{[i+1:K]}}$ for $i = 1, \cdots, K$, sequentially starting with $i = 1$. That is, the ith case corresponds to $\mathcal{S}_u = \{1, 2, \cdots, i\}$. Notice that the ith case happens if and only if $\mu \in (f'_{\sigma_i}(\tau_{\sigma_i}), f'_{\sigma_{i-1}}(\tau_{\sigma_{i-1}})]$, where we define $f'_{\sigma_0}(\tau_{\sigma_0}) = \infty$. Thus this is equivalent to considering that μ is in the intervals $(f'_{\sigma_i}(\tau_{\sigma_i}), f'_{\sigma_{i-1}}(\tau_{\sigma_{i-1}})]$ for $i = 1, 2, \cdots, K$, sequentially to decide the correct μ interval.

Algorithm 6 The fourth proposed order-based algorithm for P3.

Order $f'_k(\tau_k)$'s decreasingly, i.e., via (33), to obtain the index sequence $(\sigma_1, \cdots \sigma_K)$.
$i \leftarrow 1$, $\mu \leftarrow f'_{\sigma_i}(\tau_{\sigma_i})$.
Compute p_k from (34);
while $\sum_{k=1}^{K} p_k < P$ and $i \leq K$ **do**
  if $i < K$ **then**
    $i \leftarrow i + 1, \ \mu \leftarrow f'_{\sigma_i}(\tau_{\sigma_i})$.
    Compute p_k from (34);
  end if
end while
if $i \neq K$ **then**
  Perform **Algorithm 2** on subchannels $\{\sigma_{i+1}, \cdots, \sigma_K\}$ with the total power being $P - \sum_{k=1}^{i} \tau_{\sigma_k}$.
end if
return p_k's.

In dealing with the ith case, let $\mu = f'_{\sigma_i}(\tau_{\sigma_i})$, and the power for each subchannel is given by

$$\begin{cases} p_{\sigma_k} = \tau_{\sigma_k} & k \leq i, \\ p_{\sigma_k} = g_{\sigma_k}(\mu) & k > i. \end{cases} \tag{34}$$

Then the total power is calculated and compared with the power constraint P. If $\sum_{k=1}^{K} p_k \geq P$, none of the remaining subchannels $\sigma_i, \cdots, \sigma_K$ can reach its upper bound. Thus the optimal solution of P3 falls into this case. As for this case, subchannels $\sigma_i, \cdots, \sigma_K$ have inactive upper bounds, the bounds can be ignored and **Algorithm 2** can be used to find the optimal values of their powers. If $\sum_{k=1}^{K} p_k < P$, the increasing rate $\mu = f'_{\sigma_i}(\tau_{\sigma_i})$ is too high for all $p_{\sigma_{i+1}}, \cdots, p_{\sigma_K}$ to stay below their upper bounds. As a result, Case i is not the optimal and the next case should be considered. If the last case, Case K is considered, and still $\sum_{k=1}^{K} p_k < P$, this means $P > \sum_{i=1}^{K} \tau_k$ and all subchannels should use their maximum powers.

With the above discussions and Lemma 5, the following lemma can be proved.

Lemma 7 ***Algorithms 5*** *and* ***6*** *converge and achieve the optimal solution of P3.*

C. Application Examples

In this subsection, a few application examples are given.

Example 6: A weighted capacity maximization problem under box constraints can be formulated as follows:

$$\max_{p_1, \cdots, p_K} \quad \sum\nolimits_{k=1}^{K} w_k \log(b_k + a_k p_k)$$

$$\text{s.t.}\quad \sum\nolimits_{k=1}^{K} p_k \le P,\;\; \gamma_k \le p_k \le \tau_k, \quad i=1,\cdots,K. \quad (35)$$

Example 7: A weighted MSE minimization problem under box constraints can be written in the following form

$$\begin{aligned} \min_{p_1,\cdots,p_K} \quad & \sum\nolimits_{k=1}^{K} \frac{w_k}{b_k + a_k p_k} \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \le P,\; \gamma_k \le p_k \le \tau_k, \quad i=1,\cdots,K. \end{aligned} \quad (36)$$

Example 8: The weighted capacity maximization problem for AF MIMO relaying systems can be written in the following form

$$\begin{aligned} \min_{p_1,\cdots,p_K} \quad & \sum\nolimits_{k=1}^{K} w_k \log\left(1 - \frac{a_k b_k p_k}{1 + b_k p_k}\right) \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \le P,\; \gamma_k \le p_k \le \tau_k, \quad i=1,\cdots,K. \end{aligned} \quad (37)$$

For the problems in Examples 6-8, the proposed algorithms in Algorithms 3-6 can be used to find the optimal solutions. While in existing literature, the solutions of these problems are shown to have water-filling structure, and no systematic algorithms have been available.

IV. Several Extensions

In this section, the proposed viewpoint and algorithms are extended to several more complicated optimization problems.

A. Problems with Multiple Ascending Sum-Constraints

We first investigate the extension of P3 to have multiple ascending sum-constraints [17]. The optimization problem is posed as follows:

$$\begin{aligned} \mathrm{P4}: \max_{p_1,\cdots,p_K} \quad & \sum\nolimits_{k=1}^{K} f_k(p_k) \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{J} p_k \le P_J,\; J=1,\cdots,K \\ & \gamma_k \le p_k \le \tau_k, \quad k=1,\cdots,K, \end{aligned} \quad (38)$$

where f_k's have the same properties as in P3. The main difference to P3 is that P4 has a total of K constraints on the ascending accumulative sums, while P3 has one total sum-constraint. Thus the feasible region of P3 is no smaller than that of P4. A new approach different from that in [17] is proposed here, where the algorithms proposed for P3 in the previous section are used as building blocks.

To see the connection of P4 to P3, notice that if only the last J-th sum-constraint in (38) is considered (other sum-constraints are simply ignored), the problem is the same as P3. Denote the optimal solution of this reduced one (by considering the last sum-constraint only) as $\{p_1^\star,\cdots,p_K^\star\}$. If it satisfies all the sum-constraints in (38), it is also the optimal solution of P4. If one or more sum-constraints in (38) is violated, let $J_{\min}$ be the smallest J-value among the violated sum-constraints, that is, $\sum_{k=1}^{J} p_k^\star \le P_J$ for $J=1,\cdots,J_{\min}-1$ and $\sum_{k=1}^{J_{\min}} p_k^\star > P_{J_{\min}}$. In this case, the $J_{\min}$-th sum-constraint $\sum_{k=1}^{J_{\min}} p_k \le P_{J_{\min}}$ is a crucial one for the power allocation among the first $J_{\min}$ sub-channels, while the power allocation of the remaining subchannels can be conducted with the remaining power $P_K - P_{J_{\min}}$.

Algorithm 7 Proposed recursive nested algorithm for P4.

1: $P_0 = 0, P = P_K,\ J_{\min} = 1$;
2: Use one of **Algorithms 3-6** to find p_k's for subchannels in $\{1,2,\cdots,K\}$ considering only the sum-constraint $\sum_{k=J_{\min}}^{K} p_k \le P$ and all box constraints;
3: $n = 1$;
4: **while** $n < K$ **do**
5:   **if** $\sum_{k=J_{\min}}^{n} p_k \le P_n$ **then**
6:     $n = n + 1$;
7:   **else**
8:     Use one of **Algorithms 3-6** to find p_k's for subchannels in $\{J_{\min},\cdots,n\}$ considering only the sum-constraint $\sum_{k=J_{\min}}^{n} p_K \le P_n - P_{J_{\min}-1}$ and all box constraints;
9:     Use one of **Algorithms 3-6** to find p_k's for subchannels in $\{J_{\min}+1,\cdots,K\}$ considering only the single sum-constraint $\sum_{k=J_{\min}+1}^{K} p_k \le P_K - P_n$ and all box constraints;
10:     $J_{\min} = n + 1$;
11:   **end if**
12: **end while**
13: **return** p_k's.

Following this idea, we transform Problem P4 into the following two decomposed problems:

$$\begin{aligned} \mathrm{P4.1}: \max_{p_1,\cdots,p_{J_{\min}}} \quad & \sum\nolimits_{k=1}^{J_{\min}} f_k(p_k) \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{J_{\min}} p_k \le P_{J_{\min}}, \\ & \gamma_k \le p_k \le \tau_k, \quad k=1,\cdots,J_{\min}, \end{aligned}$$

and

$$\begin{aligned} \mathrm{P4.2}\ \max_{p_{J_{\min}+1},\cdots,p_K} \quad & \sum\nolimits_{k=J_{\min}+1}^{K} f_k(p_k) \\ \text{s.t.} \quad & \sum_{k=J_{\min}+1}^{J} p_k \le P_J - P_{J_{\min}},\; J = J_{\min}+1,\cdots,K, \\ & \gamma_k \le p_k \le \tau_k, \quad k = J_{\min},\cdots,K. \end{aligned}$$

P4.1 has only single sum-constraint over $\{p_1,\cdots,p_{J_{\min}}\}$ other than the box constraints. Thus Algorithms 3 - 6 can be used to find its optimal solution. Problem P4.2 may still have multiple sum-constraints (when $J_{\min}+1 < K$). The same procedure can be used to further decompose it until only one sum-constraint is left, then the subproblem can be solved via Algorithms 3 - 6. With the above discussion, the proposed recursive nested algorithm is presented in Algorithm 7.

Lemma 8 *Algorithm 7 finds a feasible solution of P4.*

Proof It suffices to show that the solution of the decomposed problem found in Step 8 of Algorithm 7 is feasible for each iteration. If we show that the solution found in Step 8 of the first iteration is feasible, all others can be shown to be feasible similarly. That is, we only need to show that the solution of Problem P4.1, denoted as $\{p_1^*,\cdots,p_{J_{\min}}^*\}$, satisfies all applicable conditions of Problem P4.

First it is obvious that the box constraints $\gamma_k \leq p_k^* \leq \tau_k, \quad k = 1, \cdots, J_{\min}$ are satisfied. From the construction of $p_k^\star$'s and $J_{\min}$, we can conclude that $\{p_1^\star, \cdots, p_K^\star\}$ (the solution of considering the last sum-constraint only) is also the solution of the following optimization problem:

$$\begin{aligned} \max_{p_1, \cdots, p_K} \quad & \sum\nolimits_{k=1}^{K} f_k(p_k) \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_k \leq P_K, \ \sum\nolimits_{k=1}^{J_{\min}} p_k = \sum_{k=1}^{J_{\min}} p_k^\star, \\ & \gamma_k \leq p_k \leq \tau_k, \quad k = 1, \cdots, K, \end{aligned}$$

which can be decomposed with equivalence to the following two problems:

$$\begin{aligned} \text{P4.3}: \max_{p_1, \cdots, p_{J_{\min}}} \quad & \sum\nolimits_{k=1}^{J_{\min}} f_k(p_k) \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{J_{\min}} p_k \leq \sum_{k=1}^{J_{\min}} p_k^\star, \\ & \gamma_k \leq p_k \leq \tau_k, \quad k = 1, \cdots, J_{\min}, \end{aligned}$$

and

$$\begin{aligned} \text{P4.4}: \max_{p_{J_{\min}+1}, \cdots, p_K} \quad & \sum\nolimits_{k=J_{\min}+1}^{K} f_k(p_k) \\ \text{s.t.} \quad & \sum\nolimits_{k=J_{\min}+1}^{K} p_k \leq P_K - \sum_{k=1}^{J_{\min}} p_k^\star, \\ & \gamma_k \leq p_k \leq \tau_k, \quad k = J_{\min}+1, \cdots, K. \end{aligned}$$

Thus, $\{p_1^\star, \cdots, p_{J_{\min}}^\star\}$ is the solution of P4.3.

Further from the construction of $J_{\min}$, we have $\sum_{k=1}^{J_{\min}} p_k^\star > P_{J_{\min}}$. By comparing P4.3 with P4.1 (whose solution is $\{p_1^*, \cdots, p_{J_{\min}}^*\}$), P4.3 has a larger feasible region than that of P4.1. By using the properties of f_k's and results in Lemma 5, we can conclude that $p_k^* \leq p_k^\star$. As $\sum_{k=1}^{J} p_k^\star \leq P_J$ holds for $J = 1, \cdots, J_{\min} - 1$, $\sum_{k=1}^{J} p_k^* \leq P_J$ also holds for $J = 1, \cdots, J_{\min} - 1$. ■

Generally speaking, the algorithm is not guaranteed to find the global optimal solution, but based on the decomposition approach, it is expected to be close-to-optimal. When the problem only needs to be decomposed once, it is obvious that the globally optimal solution is found.

B. Problems with Multiple Water-Levels

For MIMO-OFDM systems, some resource allocation problems aim at maximizing sum-utilities but with a certain level of fairness among subcarriers, e.g., maximizing the minimum sum-utility function among subcarriers. With the optimal diagonalizable structures, the resource allocation along the eigenchannels aligning the optimal spatial basis can be cast in the following form:

$$\begin{aligned} \text{P5}: \quad \max_{\{p_{k,j}\}} \quad & \min_{j} \sum\nolimits_{k=1}^{K} f_{k,j}(p_{k,j}) \\ \text{s.t.} \quad & \sum\nolimits_{j=1}^{J} \sum\nolimits_{k=1}^{K} p_{k,j} \leq P, \quad p_{k,j} \geq 0, \end{aligned}$$

Algorithm 8 Proposed algorithm for problems P5, P6, and P7.

```
Initialize I_{k,j} = 1 ∀k, j;
Calculate μ_j's and p_{k,j}'s by jointly solving Eqns. (41), (42), and (43);
while length(find({p_{k,j}} < 0)) > 0 do
    S_{j,inactive} = {(k, j)|p_{k,j} < 0};
    Set I_{k,j} = 0 for (k, j) ∈ S_{j,inactive};
    Calculate μ_j's and p_{k,j}'s by jointly solving Eqns. (41), (42), and (43);
end while
return p_{k,j}'s.
```

where $f_{k,j}$'s have the same properties as the f_k-functions in P2. It can be easily shown that P5 is equivalent to the following:

$$\begin{aligned} \max_{\{p_{k,j}\}} \quad & \sum\nolimits_{k=1}^{K} f_{k,j}(p_{k,j}) \\ \text{s.t.} \quad & \sum\nolimits_{j=1}^{J} \sum\nolimits_{k=1}^{K} p_{k,j} \leq P, \quad p_{k,j} \geq 0, \\ & \sum\nolimits_{k=1}^{K} f_{k,1}(p_{k,1}) = \cdots = \sum\nolimits_{k=1}^{K} f_{k,J}(p_{k,J}) = t. \end{aligned}$$

This is because at the optimum, the objective values on all subcarriers are the same. Define $\mathcal{S}_{j,active} \triangleq \{p_{k,j} > 0\}$ for $j = 1, \cdots, J$. The following lemma gives the necessary conditions of the optimal solution of P5.

Lemma 9 *The following conditions are necessary for the optimal solution of P5:*

$$\begin{cases} f'_{k_1,j}(p_{k_1,j}) = f'_{k_2,j}(p_{k_2,j}) \ \textit{for} \ k_1, k_2 \in \mathcal{S}_{j,active}, \ j = 1, \cdots J, \\ \sum_{k=1}^{K} f_{k,1}(p_{k,1}) = \cdots = \sum_{k=1}^{K} f_{k,J}(p_{k,J}) = t, \\ \sum_{j=1}^{J} \sum_{k=1}^{K} p_{k,j} = P. \end{cases} \tag{39}$$

proof The necessity of the second and third lines in (39) is obvious. To show the necessity of the first one, denote the optimal solution of P5 as $p_{k,j}^*$. It is obvious that $\{p_{1,j}^*, \cdots, p_{K,j}^*\}$ must be the optimal solution for the following reduced problem:

$$\begin{aligned} \text{P5}.j: \max_{p_{1,j}, \cdots, p_{K,j}} \quad & \sum\nolimits_{k=1}^{K} f_{k,j}(p_{k,j}) \\ \text{s.t.} \quad & \sum\nolimits_{k=1}^{K} p_{k,j} \leq \sum\nolimits_{k=1}^{K} p_{k,j}^*, \quad p_{k,j} \geq 0 \end{aligned} \tag{40}$$

since P5.j is a subproblem. From Lemma 1, the necessity of the first line of (39) is proved. ■

Thus following the viewpoint and scheme in Section II-B, we define $g_{k,j}(\cdot) \triangleq \text{Inv}[f'_{k,j}](\cdot)$ and introduce the indication operator $\mathcal{I}_{k,j}$ as: $\mathcal{I}_{k,j} = 1$ if $p_{k,j} > 0$ and $\mathcal{I}_{k,j} = 0$ if $p_{k,j} = 0$. The necessary conditions can be rewritten as

$$p_{k,j} = g_{k,j}(\mu_j)\mathcal{I}_{k,j}. \tag{41}$$

$$\sum\nolimits_{k=1}^{K} f_{k,j}(p_{k,j}) = t. \tag{42}$$

$$P = \sum\nolimits_{j=1}^{J} \sum\nolimits_{k=1}^{K} g_{k,i}(\mu_j)\mathcal{I}_{k,j}, \tag{43}$$

where μ_j is the final increasing rate (also referred to as the water level) for the jth subcarrier. **Algorithm 8** is thus proposed whose solution values satisfy the necessary conditions.

In the following, two application examples are given to demonstrate the application of Algorithm 8 for optimization problems in wireless communications.

Example 9: With the optimal diagonalization, the maximum-sum-weighted-MSE minimization problem for MIMO-OFDM systems can be formulated as

$$\begin{aligned} &\max_{\{p_{k,j}\}} \min_{j} \ -\sum\nolimits_{k=1}^{K} \frac{w_{k,j}}{b_{k,j}+a_{k,j}p_{k,j}} \\ &\text{s.t.} \quad \sum\nolimits_{j=1}^{J}\sum\nolimits_{k=1}^{K} p_{k,j} \leq P, \end{aligned} \tag{44}$$

where the term $\sum_{k=1}^{K} w_{k,j}/(b_{k,j}+a_{k,j}p_{k,j})$ is the weighted sum-weighted-MSE on the jth subcarrier. From (41),

$$p_{k,j} = g_{k,j}(\mu_j)\mathcal{I}_{k,j} = \left(\sqrt{\frac{w_{k,j}}{\mu_j a_{k,j}}} - \frac{b_{k,j}}{a_{k,j}}\right)\mathcal{I}_{k,j}. \tag{45}$$

By using (45) in Eqn. (42), we have

$$\sum_{k=1}^{K} \frac{w_{k,j}}{b_{k,j}+a_{k,j}p_{k,j}} = \sum_{k\in\mathcal{S}_{j,active}} w_{k,j} + \sum_{k\notin\mathcal{S}_{j,active}} \sqrt{\frac{a_{k,j}}{\mu_j}} = t, \tag{46}$$

from which we can solve μ_j as a function of t as follows:

$$\sqrt{\frac{1}{\mu_j}} = \frac{\sum_{k\in\mathcal{S}_{j,active}} w_{k,j}/\sqrt{a_{k,j}}}{t-\sum_{k\notin\mathcal{S}_{j,active}} w_{k,j}}. \tag{47}$$

With (47), the sum power constraint (43) can be rewritten as

$$\sum_{k=1}^{K} \frac{(\sum_{k\in\mathcal{S}_{j,active}} w_{k,j}/\sqrt{a_{k,j}})^2}{t-\sum_{k\notin\mathcal{S}_{j,active}} w_{k,j}} - \sum_{k=1}^{K}\sum_{k\in\mathcal{S}_{j,active}} \frac{b_{k,j}}{a_{k,j}} = P. \tag{48}$$

For given $\mathcal{I}_{k,j}$'s, bisection search can be used to compute t from (48). Then with (47), μ_j's can be computed and with (45), $p_{k,j}$'s can be computed.

Example 10: The maximization of the minimum weighted mutual information for MIMO-OFDM systems with optimal diagonalization can be formulated as the following:

$$\begin{aligned} &\max_{\{p_{k,j}\}} \min_{j} \ \sum\nolimits_{k=1}^{K} w_{k,j}\log|b_{k,j}+a_{k,j}p_{k,j}| \\ &\text{s.t.} \sum\nolimits_{j=1}^{J}\sum\nolimits_{k=1}^{K} p_{k,j} \leq P, \quad p_{k,j} \geq 0. \end{aligned} \tag{49}$$

This optimization plays a key role in the transceiver optimizations of MIMO-OFDM systems with nonlinear Tomlinson Harashima precoding (THP) or decision feedback equalization. Following the same logic as that for Example 9, we have via manipulating (41) and (42)

$$p_{k,j} = \left(\frac{w_{k,j}}{\mu_j} - \frac{b_{k,j}}{a_{k,j}}\right)\mathcal{I}_{k,j}. \tag{50}$$

$$\log\frac{1}{\mu_j} = \frac{t - \sum\limits_{k\notin\mathcal{S}_{j,active}} w_{k,j}\log b_{k,j} - \sum\limits_{k\in\mathcal{S}_{active}} w_{k,j}\log(a_{k,j}w_{k,j})}{\sum\limits_{k\in\mathcal{S}_{j,inactive}} w_{k,j}}. \tag{51}$$

By using the above two equations in (43), the value of t can be computed by solving the following equation:

$$\sum_{k=1}^{K}\sum_{j=1}^{J}\left(\frac{w_{k,j}}{\mu_j} - \frac{b_{k,j}}{a_{k,j}}\right)\mathcal{I}_{k,j} = P. \tag{52}$$

Because of the monotonicity of the left-hand-side of (52) with respect to t, bisection search can be used to find the value of t. The values of μ_j's can be obtained from (51), and the values of $p_{k,j}$'s can be obtained from (50).

In P5, only the zero lower bound is considered for $p_{k,j}$. It can be generalized to include box constraints. The new optimization problem is given in the following:

$$\begin{aligned} \text{P5.1}: \ &\max_{\{p_{k,j}\}} \min_{j} \sum\nolimits_{k=1}^{K} f_{k,j}(p_{k,j}) \\ &\text{s.t.} \ \sum\nolimits_{j=1}^{J}\sum\nolimits_{k=1}^{K} p_{k,j} \leq P, \ \gamma_{k,j} \leq p_{k,j} \leq \tau_{k,j} \quad p_{k,j} \geq 0. \end{aligned} \tag{53}$$

An algorithm can be designed based on the combination of **Algorithms 7** and **8**, where **Algorithm 7** provides the principle for choosing subchannels and **Algorithm 8** provides the mechanism for computing multiple water levels. Specifically, an algorithm for P5.1 can be formed by changing "one of **Algorithms 3-6**" in Steps 2,8, 9 of **Algorithm 7** to "**Algorithms 8**".

Example 11: The minimization of the maximum weighted MSE under box constraints for MIMO-OFDM systems can be cast as P5, where

$$f_{k,j}(p_{k,j}) = \frac{w_{k,j}}{b_{k,j}+a_{k,j}p_{k,j}}.$$

A solution can be found with the proposed mixed algorithm where intermediate calculations are given in (45)-(47).

Example 12: Another example is the combination of Examples 6 and 10 for the minimum capacity maximization in MIMO-OFDM systems, where

$$f_{k,j}(p_{k,j}) = w_{k,j}\log|b_{k,j}+a_{k,j}p_{k,j}|.$$

A solution can be found with the proposed mixed algorithm where intermediate calculations are given in (50)-(52).

C. Problems with Clusters

Another set of resource allocation problems have clustered water-filling structures, e.g., the power allocation in MIMO-OFDM systems under imperfect CSI. With the optimal diagonalizable structure, this type of optimization can be formulated and/or transformed in the following form [13], [20]:

$$\begin{aligned} \text{P6}: \ &\max_{\{p_{k,j}\},\{P_j\}} \ \sum\nolimits_{j=1}^{J}\sum\nolimits_{k=1}^{K} f_{k,j}(p_{k,j}, P_j) \\ &\text{s.t.} \quad \sum\nolimits_{k=1}^{K} p_{k,j} \leq P_j, \ \sum\nolimits_{j=1}^{J} P_j \leq P, \ p_{k,j} \geq 0, \end{aligned} \tag{54}$$

where $p_{k,j}$ is the power for Channel/Cluster j on Subcarrier k and $\{P_1, \cdots, P_J\}$ is a set of auxiliary variables representing the total powers over the channels across subcarriers. One important difference of P6 to the previous problems lies in the structure of $f_{k,j}$. Other than $p_{k,j}$'s, it is also a function of P_j. With respect to $p_{k,j}$ while P_j is considered fixed,

$f_{k,j}$ is assumed to have the same properties (strictly concave, increasing, and continuously differentiable) as before.

It is obvious that for given values of P_j's, the optimization problem P6 decouples into J subproblems, one for each J and all following the format of P1. Thus similar to Section II, we introduce the index operator as: $\mathcal{I}_{k,j}=1$ if $p_{k,j}>0$ and $\mathcal{I}_{k,j}=0$ if $p_{k,j}=0$. From the KKT conditions, we have

$$p_{k,j}=g_{k,j}(\mu_j,P_j)\mathcal{I}_{k,j},\ P_j=\sum_{k=1}^{K}g_{k,j}(\mu_j,P_j)\mathcal{I}_{k,j}. \quad (55)$$

The optimal value of each subproblem can be found via Algorithm 1. To solve the values of P_j's, from KKT conditions of P6, we have

$$\sum\nolimits_{k=1}^{K}\frac{\partial f_{k,j}(p_{k,j},P_j)}{\partial P_j}=\mu_j-\gamma, \quad (56)$$

$$P=\sum\nolimits_{j=1}^{J}\sum\nolimits_{k=1}^{K}g_{k,j}(\mu_j,P_j)\mathcal{I}_{k,j}. \quad (57)$$

where μ_j's are the Lagrange multipliers for the first set of constraints of P6 and λ is the multiplier for the second constraint. Based on (55)-(57), by following the framework in Section II-B, **Algorithm 8** can be used to find the solution values of P6 as well by only replacing the equations in Steps 2 and 6 to (55)-(57).

Two application examples for P6 are the weighted sum-MSE minimization [13] and the weighed sum-capacity maximization [20] for MIMO-OFDM systems under imperfect CSI. Both can be transformed into P6 with the optimal diagonalization structure. The optimal solution values can be found via **Algorithm 8**, where more details are omitted here and interested readers are referred to [13] and [20].

D. Problems with Clusters and Multiple Water Levels

A combination of P5 and P6 can be formulated as follows:

$$\begin{aligned}\text{P7}:\quad &\max_{\{p_{k,j}\},\{P_j\}}\ \min_j \sum\nolimits_{k=1}^{K}f_{k,j}(p_{k,j},P_j)\\ \text{s.t.}\quad &\sum\nolimits_{k=1}^{K}p_{k,j}\le P_j,\ \sum\nolimits_{j=1}^{J}P_j\le P,\ p_{k,j}\ge 0,\end{aligned}$$

where $f_{k,j}$'s have the same properties as in P6. The problem models the optimization of MIMO-OFDM systems under imperfect CSI with consideration of fairness. By following the derivations in Sections IV-B and IV-C, we have, from the KKT conditions,

$$p_{k,j}=g_{k,j}(\mu_j,P_j)\mathcal{I}_{k,j}. \quad (58)$$

$$P_j=\sum_{k=1}^{K}g_{k,j}(\mu_j,P_j)\mathcal{I}_{k,j}. \quad (59)$$

$$\sum\nolimits_{k=1}^{K}f_{k,j}(p_{k,j},P_j)=t. \quad (60)$$

$$P=\sum\nolimits_{j=1}^{J}\sum\nolimits_{k=1}^{K}g_{k,j}(\mu_j,P_j)\mathcal{I}_{k,j}. \quad (61)$$

Algorithm 8 can be used to find the solution values of P6 as well by only replacing the equations in Steps 2 and 6 with (58)-(61).

Example 13: With optimal diagonalization structure, the problem of the minimum weighted sum-capacity maximization for MIMO-OFDM systems under imperfect CSI can also be formulated as P7 where

$$f_{k,j}(p_{k,j},P_j)=w_{k,j}\log\left(1+\frac{a_{k,j}p_{k,j}}{\sigma_{e_j}^2P_j+\sigma_n^2}\right).$$

By jointly solving (58) and (61), the values of $p_{k,j}$'s can be obtained for any given $\mathcal{I}_{k,j}$. Thus Algorithm 8 can be used to find the solution values. This algorithm can also be applied to the minimization of the maximum weighted sum-MSE under imperfect CSI.

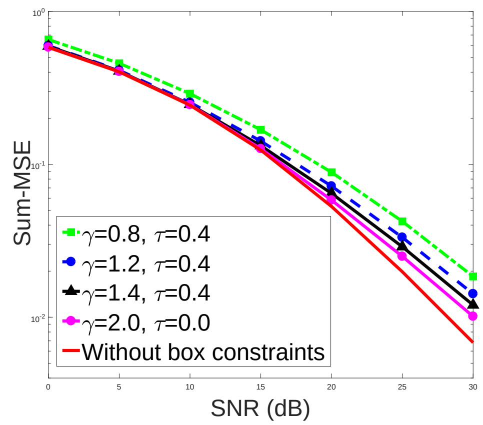

Fig. 1. The MSEs under different box constraints for single-user MIMO-OFDM system with 4 antennas at transceiver nodes.

V. Numerical Results

Due to space limitation, only the most general and representative cases are simulated and shown. MIMO-OFDM systems are adopted as the most complicated case considered in this work. Further, optimization related to MSE minimization is chosen since it is more complicated and interesting than those for capacity maximization. For high signal-to-noise-ratio (SNR) regime, the solutions for capacity maximization approach uniform power allocation and as a result, the effect of box constraints becomes trivial.

It is assumed that the source and destination are each equipped with 4 antennas. A multi-path channel with 7 paths is simulated. Specifically, each channel tap is generated according to a Gaussian distribution. Further, the time-domain decaying factor of the channel taps is 0.5, i.e., the covariance of the lth tap is $1/2^{l-1}$ of that of the first tap and the sum variances of all the taps is normalized to one. The number of frequency-domain subcarriers is $J=256$. Each point in the figures is obtained by an average over 10^3 independent channel realizations. The maximum transmit power is denoted by P and the SNR is defined as $P/(N\sigma_n^2)$.

We simulate the case with perfect CSI. The sum-MSE minimization can be formulated as Problem P3, where the following box constraints are used:

$$\gamma P/(4N)\le p_{k,j}\le \tau P/(4N). \quad (62)$$

The parameters γ and τ are introduced to adjust the bounds of the power allocations to exhibit the effect of box constraints.

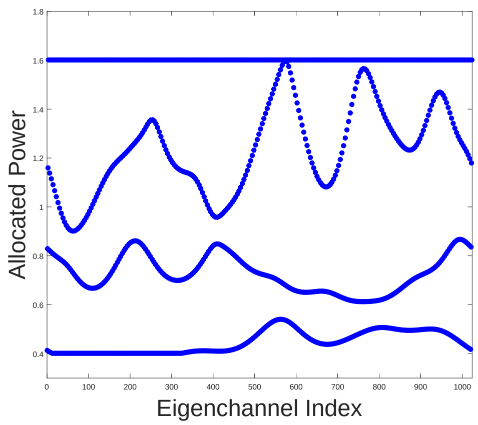

Fig. 2. The power allocation for one channel realization at the SNR of 20 dB with $\gamma = 0.4$ and $\tau = 1.6$.

The sum-MSE is shown in Fig. 1, for different SNR values, where it can be observed that the box constraints can control the fairness among different subcarriers.

In Fig. 2, we show the power allocation results across the 1024 eigen-channels for one channel realization at the SNR of 20 dB with different values of γ and τ. It can be seen that box constraints are of practical importance in the communication design as the channel qualities of different eigenchannels fluctuate significantly. Further based on the proposed algorithm, both the lower bounds and upper bounds can be met.

VI. Conclusions

Optimization problems with water-filling solutions widely arise and are fundamental for the resource allocation in wireless communications and networking. To find the solution values, practical and efficient algorithms should be carefully designed. While existing algorithms are case-by-case, sparse, and many with overly-complicated algorithm designs, in this work, a new viewpoint for such optimization problems has been proposed by understanding the power allocation procedure dynamically and considering the changes of the increasing rates on each subchannel. With this viewpoint and rigorous analysis of the solution structure, a comprehensive framework for algorithm designs has been presented in this paper. Five different kinds of optimization problems have been studied sequentially according to their complexities and efficient algorithms have been proposed. Based on our results, it can be concluded that the various algorithm designs share common fundamentals. We also expect that the proposed design logic and algorithms can be used to resolve new optimization problems in future wireless systems.

References

[1] T. Cover and J. Thomas, *Elements of Information Theory, 2nd Edition.* Wiley&Sons, 2006.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[3] M. Dai, S. Zhang, B. Chen, X. Lin, and H. Wang, "A refined convergence condition for iterative waterfilling algorithm," *IEEE Commun. Lett.,* vol. 18, no. 2, Feb. 2014.

[4] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of water-filling solutions," *IEEE Trans. Signal Process.,* vol. 53, no. 2, pp. 686–695, Feb. 2005.

[5] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, pp. 2381–2401, Sept. 2003.

[6] O. Ozel, K. Shahzad, and S. Ulukus, "Optimal energy allocation for energy harvesting transmitter with hybrid energy storage and processing cost," *IEEE Trans. Signal Process.,* vol. 62, no. 12, pp. 3232–3245, June 15, 2014.

[7] L. Lai and H. El Gamal, "The water-filling game in fading multiple-access channels," *IEEE Trans. Inf. Theory,* vol. 54, no. 5, pp. 2110-2122, May 2008.

[8] H. Moon, "Waterfilling power allocation at high SNR regimes," *IEEE Trans. Commun.,* vol. 59, no. 3, pp. 708–715, March 2011.

[9] C. Xing, Z. Fei, Y. Zhou, Z. Pan, and H. Wang, "Transceiver designs with matrix-version water-filling architecture under mixed power constraints," *Science China-Information Sciences,* vol. 59, no. 10, pp. 1–13, Oct. 2016.

[10] O. Popescu, D. C. Popescu, and C. Rose, "Simultaneous water filling in mutually interfering systems," *IEEE Trans. Wireless Commun.,* vol. 6, no. 3, pp. 1102–1113, March 2007.

[11] S. Khakurel, C. Leung, and T. Le-Ngoc, "A generalized water-filling algorithm with linear complexity and finite convergence time," *IEEE Commun. Lett.,* vol. 3, no. 2, April 2014.

[12] D. Hoang and R. A. Iltis, "Noncooperative eigencoding for MIMO ad hoc networks," *IEEE Trans. Signal Process.,* vol. 56, no. 2, pp. 865–869, Feb. 2008.

[13] C. Xing, D. Li, S. Ma, Z. Fei, and J. Kuang, "Robust transceiver designs for MIMO-OFDM systems based on cluster water-filling," *IEEE Commun. Lett.*, vol. 17, no. 7, pp. 1451–1454, July 2013

[14] X. Liang, B. Wu, P.-H. Ho, F. Luo, and L. Pan, "Fast water-filling for agile power allocation in multi-channel wireless communications," *IEEE Commun. Lett.,* vol. 16, no. 8, pp. 1212–1215, Aug. 2012.

[15] C. Xing, F. Gao, and Y. Zhou, "A framework for transceiver designs for multi-hop communications with covariance shaping constraints," *IEEE Trans. Signal Process.*, vol. 63, no. 15, pp. 3930–3945, Aug. 2015.

[16] F. Gao, T. Cui, and A. Nallanathan, "Optimal training design for channel estimation in decode-and-forward relay networks with individual and total power constraints," *IEEE Trans. Signal Process.,* vol. 56, no. 12, pp. 5937–5949, Dec. 2008.

[17] A. A. D'Amicao, L. Sanguinetti, and D. P. Palomar, "Convex separable problems with linear constraints in signal processing and communications," *IEEE Trans. Signal Process.,* vol. 62, no. 22, pp. 6045–6058, Nov. 2014.

[18] P. He, L. Zhao, S. Zhou and Z. Niu, "Water-filling: A geometric approach and its application to solve generalized radio resource allocatoon problems," *IEEE Trans. Wireless Commun.,* vol. 12, no. 7, pp. 3637–3647, July 2013.

[19] P. He, L. Zhao, S. Zhou, and Z. Niu, "Recursive waterfilling for wireless links with energy harvesting transmitters," *IEEE Trans. Veh. Technol.,* vol. 63, no, 3, pp. 1232–1241, March 2014.

[20] S. Guo, C. Xing, Z. Fei and D. Li, "Robust capacity maximization transceiver design for MIMO OFDM systems," *SCIENCE CHINA Information Sciences* 59(6): 062301:1-062301:11 (2016)

[21] C. Xing, Y. Ma, Y. Zhou, and F. Gao, "Transceiver optimization for multi-hop communications with per-antenna power constraints," *IEEE Trans. Signal Process.,* vol. 64, no. 6, pp. 1519–1534, March 2016.

[22] E. Björnson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbances," *IEEE Trans. Signal Process.,* vol. 58, no. 3, pp. 1807–1820, March 2010.

[23] J. Dai, C. Chang, W. Xu, and Z. Ye, "Linear precoder optimization for MIMO systems with joint power constraints," *IEEE Trans. Commun.* vol. 60, no. 8, pp. 2240–2254, Aug. 2012.

[24] C. Xing, S. Ma, Z. Fei, Y.-C. Wu, and H. V. Poor, "A general robust linear transceiver design for amplify-and-forward multi-hop MIMO relaying systems," *IEEE Trans. Signal Process.,* vol. 61, pp. 1196–1209, Mar. 2013.