ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-6327-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICASSP49357.2023.10096377

# ENHANCING THE EFFICIENCY OF WMMSE AND FP FOR BEAMFORMING BY MINORIZATION-MAXIMIZATION

*Zepeng Zhang*[†], *Ziping Zhao*[†], *and Kaiming Shen*[‡]

[†]School of Information Science and Technology, ShanghaiTech University, Shanghai, China
[‡]School of Science and Engineering, The Chinese University of Hong Kong (Shenzhen), China

## ABSTRACT

Weighted minimum mean squared error (WMMSE) and fractional programming (FP) constitute two common approaches to the weighted sum-rate maximization in communication system design. One subtle issue with WMMSE and FP lies in the tuning of a Lagrange multiplier for the power constraint when it comes to the multi-antenna transmission. To obtain the optimal Lagrange multiplier, we must repeatedly inverse an $M \times M$ matrix, where $M$ is the number of transmit antennas, which incurs considerable complexity. To address the above issue, this work explores the connection of WMMSE and FP to minorization-maximization (MM), thereby modifying the two methods to get rid of the Lagrange multiplier. The proposed algorithm enables a parameter-free iterative optimization of the beamforming vectors with the power constraint enforced automatically. Numerical results demonstrate the faster convergence of the proposed beamforming method as compared to the conventional WMMSE and FP methods.

***Index Terms***— Weighted sum-rate maximization, weighted minimum mean squared error (WMMSE), fractional programming (FP), minorization-maximization (MM).

## 1. INTRODUCTION

Weighted sum-rate (WSR) maximization plays a central role in a wide variety of communication system design tasks [1], a typical case of which is to find the optimal beamforming vectors for the multi-antenna channels [2–4]. This problem proves to be NP-hard [5, 6]; the state of the art is to attain a stationary point solution via the weighted minimum mean squared error (WMMSE) algorithm or fractional programming (FP). This work shows that the computational efficiency of WMMSE and FP can be significantly improved by utilizing their connection to minorization-maximization (MM) [7].

As two common approaches to the WSR beamforming problem, WMMSE and FP are driven by completely diverse motivations. WMMSE stems from a well-known result in the signal processing field: maximizing the signal-to-interference-plus-noise ratio (SINR) is equivalent to minimizing the mean squared error (MSE) of the received signal. The idea of casting the WSR maximization problem into the weighted sum MSE minimization problem to facilitate the problem solving is first proposed in [2] for the multiple-input single-out (MISO) channels and then is extended by [3] to the multiple-input multiple-output (MIMO) case. A recent progress in WMMSE aims to reduce the computational complexity by using the range space of channels under some certain conditions [8]. In contrast, the FP method works towards a mathematical goal—it reformulates a fractional optimization problem with one or more ratios so as to decouple every ratio term. The classic methods, i.e., Dinkelbach's algorithm and Charnes-Cooper algorithm [9], have achieved this goal for a single ratio. Thus, the use of the classic FP techniques in communication system design is typically restricted to the single-ratio scenario such as the efficiency maximization [10, 11], but the WSR problem does not fall in this category. In [4], a multi-ratio FP technique called the quadratic transform is developed to coordinate multiple links (with multiple SINRs) in wireless networks. The transforms in multi-ratio FP increase the dimension of the optimization variable, i.e., leading to multi-block optimization problems with introduced auxiliary variables. Then the beamforming variable and some auxiliary variables are optimized in a block coordinate ascent (BCA) fashion [12]. The capability of dealing with multiple ratios enables the extensive applications of FP in communication system design, e.g., [13–16].

Nevertheless, it turns out that WMMSE and FP are closely related to each other, and are both akin to MM. As shown in [17, 18], WMMSE can be recognized as a special case of FP [4], and moreover, FP can be recognized as a special case of the more flexible MM method. One main result of this paper is to establish the above connection using a novel constructive argument, which is inspired by the MM algorithms for WSR maximization proposed in [19, 20]. In return, the MM interpretation leads us to a novel way of enhancing the efficiency of WMMSE and FP. Actually, a subtle issue with WMMSE and FP is to decide a Lagrange multiplier for each transmit beamformer—which is introduced to account for the transmit power constraint. According to the conventional WMMSE and FP, we must repeatedly inverse an $M \times M$ matrix in order to obtain the optimal Lagrange multi-

This work was supported in part by the National Nature Science Foundation of China (NSFC) under Grant 62001295, in part by the Shanghai Sailing Program under Grant 20YF1430800, and in part by the NSFC 62001411.

plier, where $M$ is the number of transmit antennas. There are some learning-based methods designed to avoid the matrix inversion [21], while no optimization-based approach has been presented to achieve such a goal. The other main result of the paper is to modify the existing WMMSE and FP methods to get rid of the Lagrange multiplier by leveraging their connections to MM [20]. Like WMMSE and FP, the proposed method still optimizes the beamforming variable and some auxiliary variables in a BCA fashion, only that the tuning of the Lagrange multiplier for the power constraint is no longer required. The algorithm yields a provable convergence to a feasible stationary point satisfying the constraint.

The rest of the paper is organized as follows. Section 2 describes the WSR beamforming problem. For ease of notation, we focus on a single cell with the MISO channels, but our results can be readily extended to the fully general multi-cell MIMO case. Section 3 examines the connection of FP and WMMSE to MM. Section 4 shows the proposed beamforming algorithm. Section 5 presents the simulation results. Finally, Section 6 concludes the entire paper.

## 2. WSR MAXIMIZATION IN MISO SYSTEMS

Consider a single-cell MISO downlink network with $K$ user terminals, where the base station (BS) has $M$ antennas and each user terminal has one antenna. Let $\mathbf{h}_k \in \mathbb{C}^{1\times M}$ be the channel from the BS to the $k$th user terminal. Let $\mathbf{w} \in \mathbb{C}^M$ be the transmit beamforming vector for the downlink transmission to the $k$th user terminal. We write $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_K]$ as a shorthand for the beamforming variables. Thus, the SINR of each user $k$ can be computed as

$$\mathsf{SINR}_k = \frac{\left|\mathbf{h}_k^H \mathbf{w}_k\right|^2}{\sum_{j=1,j\neq k}^K \left|\mathbf{h}_k^H \mathbf{w}_j\right|^2 + \sigma_k^2}, \tag{1}$$

where $\sigma^2$ is the additive background noise power. In particular, a sum power constraint $\sum_{k=1}^K \|\mathbf{w}_k\|^2 \le P$ is posed at the transmitter side. We use $\mathcal{W}$ to denote the feasible set of $\mathbf{W}$ under the power constraint $P$, i.e., $\mathcal{W} = \left\{\mathbf{W} : \sum_{k=1}^K \|\mathbf{w}_k\|^2 \le P\right\}$. Moreover, a nonnegative rate weight $\alpha_k \ge 0$ is assigned to each user $k$ to reflect its priority. The WSR maximization problem across the $K$ downlink users can now be formulated as

$$\underset{\mathbf{W}\in\mathcal{W}}{\mathsf{maximize}} \quad f_{\mathsf{WSR}}(\mathbf{W}) \triangleq \sum_{k=1}^K \alpha_k \log\left(1 + \mathsf{SINR}_k\right). \tag{2}$$

## 3. MM INTERPRETATION OF WMMSE AND FP

### 3.1. Review of WMMSE and FP

Because WMMSE is a special case of FP as shown in [17, Section VI-A], it suffices to show how FP works. The FP method [4] is based on the following two results:

**Theorem 1** ( [17]). *Given ratios $\frac{A_k(\mathbf{x})}{B_k(\mathbf{x})}$ with $A_k(\mathbf{x}) \ge 0$ and $B_k(\mathbf{x}) > 0$ for $k = 1, \ldots, K$, the following problem:*

$$\underset{\mathbf{x}\in\mathcal{X}}{\mathsf{maximize}} \quad \sum_{k=1}^K \alpha_k \log\left(1 + \frac{A_k(\mathbf{x})}{B_k(\mathbf{x})}\right) \tag{3}$$

*is equivalent to*

$$\underset{\mathbf{x}\in\mathcal{X},\boldsymbol{\gamma}}{\mathsf{maximize}} \sum_{k=1}^K \omega_k\Big(\log(1+\gamma_k) - \gamma_k + \frac{(1+\gamma_k)\,A_k(\mathbf{x})}{A_k(\mathbf{x}) + B_k(\mathbf{x})}\Big), \tag{4}$$

*in the sense that they attain identical optimal solution, where $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_K]$.*

**Theorem 2** ( [4]). *Given nondecreasing functions $f_k$ and ratios $\frac{|A_k(\mathbf{x})|^2}{B_k(\mathbf{x})}$ with $A_k(\mathbf{x}) \in \mathbb{C}$ and $B_k(\mathbf{x}) > 0$ for $k = 1, \ldots, K$, the following problem:*

$$\underset{\mathbf{x}\in\mathcal{X}}{\mathsf{maximize}} \quad \sum_{k=1}^K f_k\left(\frac{|A_k(\mathbf{x})|^2}{B_k(\mathbf{x})}\right) \tag{5}$$

*is equivalent to*

$$\underset{\mathbf{x}\in\mathcal{X},\mathbf{y}}{\mathsf{maximize}} \quad \sum_{k=1}^K f_k\left(2\mathrm{Re}\left(y_k^* A_k(\mathbf{x})\right) - |y_k|^2 B_k(\mathbf{x})\right), \tag{6}$$

*in the sense that they attain identical optimal solution, where $\mathbf{y} = [y_1, \ldots, y_K]$.*

Applying the Lagrangian dual transform in Theorem 1 to Prob. (2), we reformulate the problem as

$$\underset{\mathbf{W}\in\mathcal{W},\boldsymbol{\gamma}}{\mathsf{maximize}} \quad \sum_{k=1}^K \alpha_k\Bigg(\log(1+\gamma_k) - \gamma_k + \frac{(1+\gamma_k)|\mathbf{h}_k^H\mathbf{w}_k|^2}{\sum_{j=1}^K |\mathbf{h}_k^H\mathbf{w}_j|^2 + \sigma_k^2}\Bigg). \tag{7}$$

Prob. (7) is convex in $\gamma$ when the other variable is held fixed. Given $\underline{\mathbf{W}}$[1], each $\gamma_k$, $k = 1, \ldots, K$, can be optimally determined as $\gamma_k^\star = \underline{\mathsf{SINR}}_k$. The ratios in Prob. (7) can be further decoupled using the quadratic transform in Theorem 2, and thus we arrive at a further reformulation:

$$\underset{\mathbf{W}\in\mathcal{W},\boldsymbol{\gamma},\mathbf{y}}{\mathsf{maximize}} \sum_{k=1}^K\Big(2\mathrm{Re}\big(y_k^*\sqrt{\alpha_k(1+\gamma_k)}\mathbf{h}_k^H\mathbf{w}_k\big) - |y_k|^2 \cdot \Big(\sigma_k^2 + \sum_{j=1}^K\left|\mathbf{h}_k^H\mathbf{w}_j\right|^2\Big) + \omega_k\big(\log(1+\gamma_k) - \gamma_k\big)\Big). \tag{8}$$

Thus, aside from the original beamforming variable $\mathbf{W}$, two auxiliary variables $\boldsymbol{\gamma}$ and $\mathbf{y}$ are introduced by the Lagrangian dual transform and the quadratic transform, respectively. The FP method is to optimize the above three variables iteratively, as shown below:

[1] We use $\underline{\mathbf{x}}$ denotes the given variables (e.g., $\underline{\mathbf{w}}_k$) or computed with given variables (e.g., $\underline{\mathsf{SINR}}_k$) with most recently updated value.

FP Approach to WSR Beamforming

(Step 1) $\gamma_k^\star = \underline{\mathsf{SINR}}_k$,

(Step 2) $y_k^\star = \dfrac{\sqrt{\omega_k\left(1+\underline{\gamma}_k\right)}\mathbf{h}_k^H\underline{\mathbf{w}}_k}{\sum_{j=1}^K\left|\mathbf{h}_k^H\underline{\mathbf{w}}_j\right|^2+\sigma_k^2}$,

(Step 3) $\mathbf{w}_k^\star = \left(\sum_{j=1}^K|\underline{y}_j|^2\mathbf{h}_j\mathbf{h}_j^H+\mu^\star\mathbf{I}\right)^\dagger\sqrt{\omega_k(1+\underline{\gamma}_k)}\underline{y}_k\mathbf{h}_k$.

Go back to Step 1 till convergence

We remark that the quadratic transform can be applied in different ways to Prob. (7). For instance, we could have excluded the term $\alpha_k(1+\gamma_k)$ from the ratio term by treating it as a weight of the ratio, then would arrive at a different problem:

$$\begin{aligned}\underset{\mathbf{W}\in\mathcal{W},\boldsymbol{\gamma},\mathbf{y}}{\text{maximize}}\ &\sum_{k=1}^K\alpha_k\left(1+\gamma_k\right)\Big(2\mathrm{Re}\big(y_k^*\mathbf{h}_k^H\mathbf{w}_k\big)-|y_k|^2\cdot\\ &\big(\sigma_k^2+\sum_{j=1}^K\left|\mathbf{h}_k^H\mathbf{w}_j\right|^2\big)+\omega_k\big(\log\left(1+\gamma_k\right)-\gamma_k\big)\Big).\end{aligned}\tag{9}$$

As shown in [17], optimizing the three variables $\mathbf{W}$, $\boldsymbol{\gamma}$, and $\mathbf{y}$ iteratively gives rise to the WMMSE algorithm. Thus, WMMSE can be viewed as a special case of FP. The pros and cons of the different patterns of ratio decoupling has been discussed in [17].

Another remark we wish to make is about the Lagrange multiplier $\mu$ nested in the update of $\mathbf{w}_k$ at Step 3 of the above algorithm. As shown in [17], $\mu$ accounts for the sum power constraint, whose optimal value can be determined according to the complementary slackness

$$\mu^\star=\min\Big\{\mu\geq 0:\sum_{k=1}^K\left\|\mathbf{w}_k\left(\mu\right)\right\|_2^2\leq P\Big\}.\tag{10}$$

However, the search for the optimal $\mu^\star$ entails computing the $M\times M$ matrix pseudo-inverse $\left(\sum_{j=1}^K|\underline{y}_j|^2\mathbf{h}_j\mathbf{h}_j^H+\mu^\star\mathbf{I}\right)^\dagger$ repeatedly with respect to different possible values of $\mu^\star$, which can be time costly in practice. Both WMMSE and FP are faced with this issue; this work aims to get rid of the Lagrange multiplier $\mu$ by MM while still satisfying the sum power constraint on $\mathbf{W}$. Before proceeding to this main result, we first explore the connection between FP and MM.

### 3.2. Connection to MM

Since WMMSE is a special case of FP as shown in the previous subsection, it suffices to consider how FP is connected to MM. In contrast to [17] that links FP to MM in a reverse engineering fashion, this work provides a constructive analysis. We begin with two lemmas:

**Lemma 3** ( [20]). *The $\log(z)$ with $z\in\mathbb{R}_+$ is minorized at $\underline{z}$ as follows:*

$$\log(z)\geq\log(\underline{z})+1-\frac{\underline{z}}{z}.$$

**Lemma 4** ( [20]). *Given a nondecreasing function $f$, $f\big(\frac{|z_1|^2}{z_2}\big)$ with $z_1\in\mathbb{C}$ and $z_2\in\mathbb{R}_+$ is minorized at $(\underline{z_1},\underline{z_2})$ as follows:*

$$f\Big(\frac{|z_1|^2}{z_2}\Big)\geq f\Big(2\mathrm{Re}\big(\frac{\underline{z_1}^*}{\underline{z_2}}z_1\big)-\frac{|\underline{z_1}|^2}{\underline{z_2}^2}z_2\Big).$$

Based on Lemmas 3 and 4, an MM method is proposed for WSR maximization in both MISO and MIMO cases in [20]. Then, we show that the Lagrangian dual transform and the quadratic transform boil down to constructing surrogate functions for MM, as stated in the following two theorems.

**Theorem 5.** *Consider the Lagrangian dual transform in Theorem 1, if we consider the optimal $\boldsymbol{\gamma}$ as a function of $\mathbf{x}$ and substitute them into the objective in Eq.* (4)*, then the resulting function is a surrogate function of the objective in Prob.* (3) *constructed based on Lemma 3.*

*Proof.* Given $\underline{\mathbf{x}}$, $\gamma_k$ can be optimally determined as $\gamma_k^\star=\frac{A_k(\underline{\mathbf{x}})}{B_k(\underline{\mathbf{x}})}$. By substituting $\boldsymbol{\gamma}^\star$ to the objective in (4), we have

$$g\left(\mathbf{x},\boldsymbol{\gamma}^\star\right)=\sum_{k=1}^K\omega_k\Big(\log\big(1+\frac{A_k(\underline{\mathbf{x}})}{B_k(\underline{\mathbf{x}})}\big)+1-\frac{1+\frac{A_k(\underline{\mathbf{x}})}{B_k(\underline{\mathbf{x}})}}{1+\frac{A_k(\mathbf{x})}{B_k(\mathbf{x})}}\Big).$$

It can be readily verified that $g\left(\mathbf{x},\boldsymbol{\gamma}^\star\right)$ is a minorizing function of the objective function in (3) at $\underline{\mathbf{x}}$, which can be constructed based on Lemma 3 by taking $z=1+\frac{A_k(\mathbf{x})}{B_k(\mathbf{x})}$. □

**Theorem 6.** *Consider the quadratic transform in Theorem 2, if we consider the optimal $\mathbf{y}$ as a function of $\mathbf{x}$ and substitute them into the objective in Eq.* (6)*, then the resulting function is a surrogate function of the objective in Prob.* (5) *constructed based on Lemma 4.*

*Proof.* With $\mathbf{x}$ being held fixed, $y_k$ can be optimally determined as $y_k^\star=\frac{A_k(\underline{\mathbf{x}})}{B_k(\underline{\mathbf{x}})}$, $k=1,\ldots,K$. By substituting $\mathbf{y}^\star$ to the objective in (6), we have

$$g\left(\mathbf{x},\mathbf{y}^\star\right)=\sum_{k=1}^K f_k\left(2\mathrm{Re}\Big(\frac{A_k^*(\underline{\mathbf{x}})}{B_k(\underline{\mathbf{x}})}A_k(\mathbf{x})\Big)-\Big|\frac{A_k(\underline{\mathbf{x}})}{B_k(\underline{\mathbf{x}})}\Big|^2B_k(\mathbf{x})\right).$$

It can be verified that $g\left(\mathbf{x},\mathbf{y}^\star\right)$ is a minorizing function of the objective in Prob. (5) at $\underline{\mathbf{x}}$, which can be constructed based on Lemma 4 with $z_1=A_k(\mathbf{x})$ and $z_2=B_k(\mathbf{x})$. □

The above results demonstrate that the update of auxiliary variables in FP can be recognized as generating the surrogate function in MM based on Lemma 3 and Lemma 4.

## 4. PROPOSED BEAMFORMING METHOD

The main issue with the conventional FP method is that it is inefficient to repeatedly compute the matrix pseudo-inverse in order to obtain the optimal $\mu$. We now propose substituting this matrix with a diagonal matrix for which the pseudo-inverse can be immediately obtained. The main idea is to use a novel MM technique as specified in the following lemma.

$$\underset{\mathbf{W}\in\mathcal{W},\boldsymbol{\gamma},\mathbf{y},\mathbf{T}\in\mathcal{W}}{\text{maximize}} \quad \sum_{k=1}^{K}\Big(2\mathrm{Re}\big(y_k^*\sqrt{\omega_k(1+\gamma_k)}\mathbf{h}_k^H\mathbf{w}_k\big) - |y_k|^2\sigma_k^2 - |y_k|^2\sum_{j=1}^{K}\Big(\mathbf{w}_k^H\,\|\mathbf{h}_j\|_2^2\,\mathbf{w}_k + 2\mathrm{Re}\big(\mathbf{w}_k^H(\mathbf{h}_j\mathbf{h}_j^H - \|\mathbf{h}_j\|_2^2\,\mathbf{I})\mathbf{t}_k\big) + \mathbf{t}_k^H(\|\mathbf{h}_j\|_2^2\,\mathbf{I} - \mathbf{h}_j\mathbf{h}_j^H)\mathbf{t}_k\Big) + \omega_k\left(\log\left(1+\gamma_k\right)-\gamma_k\right)\Big). \tag{11}$$

**Lemma 7** ( [20]). *Let $\mathbf{L},\mathbf{M}\in\mathbb{H}^n$ such that $\mathbf{M}\succeq\mathbf{L}$. The function $\mathbf{x}^H\mathbf{L}\mathbf{x}$ with $\mathbf{x}\in\mathbb{C}^n$ is majorized at $\underline{\mathbf{x}}$ as follows:*

$$\mathbf{x}^H\mathbf{L}\mathbf{x} \le \mathbf{x}^H\mathbf{M}\mathbf{x} + 2\mathrm{Re}(\mathbf{x}^H(\mathbf{L}-\mathbf{M})\underline{\mathbf{x}}) + \underline{\mathbf{x}}^H(\mathbf{M}-\mathbf{L})\underline{\mathbf{x}}.$$

In Section 3.2, we have shown that the updates of auxiliary variables in FP can be seen as procedures for generating surrogate functions. From an alternative view, these transforms are specifications for parameterizing the intermediate constants in MM. With such understanding, we introduce a novel equivalent transform based on reparameterizing $\underline{\mathbf{x}}$ in Lemma 7 as an auxiliary variable. We point out that with the close relationship between the surrogate function construction and the equivalent transform, new transforms can be derived.

**Theorem 8.** *Let $\mathbf{L},\mathbf{M}\in\mathbb{H}^n$ such that $\mathbf{M}\succeq\mathbf{L}$. Problem*

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{minimize}} \;\; \mathbf{x}^H\mathbf{L}\mathbf{x}$$

*is equivalent to*

$$\underset{\mathbf{x}\in\mathcal{X},\mathbf{t}\in\mathcal{X}}{\text{minimize}} \;\; \mathbf{x}^H\mathbf{M}\mathbf{x} + 2\mathrm{Re}(\mathbf{x}^H(\mathbf{L}-\mathbf{M})\mathbf{t}) + \mathbf{t}^H(\mathbf{M}-\mathbf{L})\mathbf{t},$$

*in the sense that they attain identical optimal solution.*

*Proof.* The auxiliary variable $\mathbf{t}$ can be optimally determined by setting the derivative of the transformed objective w.r.t. $\mathbf{t}$ to $\mathbf{0}$, by substituting which back to the transformed objective, we can get $\mathbf{x}^H\mathbf{L}\mathbf{x}$. □

We now apply Theorem 8 to Prob. (8) to further recast the problem into (11) displayed at the top of the page where $\mathbf{T}=[\mathbf{t}_1,\ldots,\mathbf{t}_K]$. Denoting

$$\mathbf{q}_k = \frac{\underline{y}_k\sqrt{\omega_k(1+\underline{\gamma}_k)}\mathbf{h}_k - \sum_{j=1}^{K}|\underline{y}_j|^2(\mathbf{h}_j\mathbf{h}_j^H - \|\mathbf{h}_j\|_2^2\,\mathbf{I})\underline{\mathbf{t}}_k}{\sum_{j=1}^{K}|\underline{y}_j|^2\,\|\mathbf{h}_j\|_2^2},$$

the update rule of a novel FP+ method is given as follows:

FP+ Approach to WSR Beamforming

(Step 1) $\gamma_k^\star = \underline{\mathsf{SINR}}_k$,

(Step 2) $y_k^\star = \dfrac{\sqrt{\omega_k\left(1+\underline{\gamma}_k\right)}\mathbf{h}_k^H\underline{\mathbf{w}}_k}{\sum_{j=1}^{K}\left|\mathbf{h}_k^H\underline{\mathbf{w}}_j\right|^2+\sigma_k^2}$,

(Step 3) $\mathbf{t}_k^\star = \underline{\mathbf{w}}_k$,

(Step 4) $\mathbf{w}_k^\star = \mathbf{q}_k\min\left\{\sqrt{\dfrac{P}{\sum_{k=1}^{K}\|\mathbf{q}_k\|_2^2}},1\right\}$.

Go back to Step 1 till convergence

Observe that $\mathbf{W}$ can now be efficiently determined without searching for the optimal $\mu$.

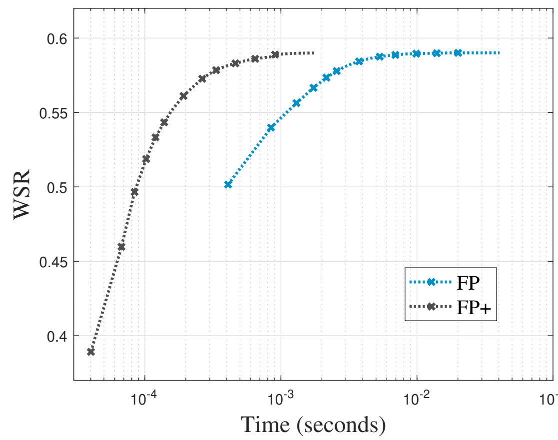


**Fig. 1**. Performance of algorithms for WSR maximization.

## 5. EXPERIMENTS

In this section, we provide numerical experiments to corroborate our theoretical results, which are performed in MATLAB on a personal computer with a 3.3 GHz Intel Xeon W CPU. Under a three-dimensional Cartesian coordinate system, we consider a multi-user MISO system, where a BS located at $(0,0,10)$m communicates with $K$ users that are randomly distributed in a circle centered at $(d,30,0)$m with radius of 10m. We adopt the Rayleigh fading model $\mathbf{h}_k=\mathbf{h}_k^{\mathsf{R}}\sqrt{\kappa(d)}$, where $\mathbf{h}_k^{\mathsf{R}}\sim\mathcal{CN}(0,1)$ and $\kappa(d)=T_0(d)^{-\varrho}$ with $T_0=-30$dB and $\varrho=3.67$. We consider the noise power spectrum density of $-169$dBm/Hz and the transmission bandwidth of 240kHz. Besides, we set $P=0$dBm $d=200$m, $K=4$, $M=4$, and $\sigma_1^2=\cdots=\sigma_K^2=1$. All the simulation curves are averaged over 100 independent channel realizations.

We compare FP and FP+ in terms of computation time are depicted in Fig. 1. Observe that although FP+ need more iterations to converge, they have faster convergence in terms of CPU times. This result suggests that FP+ requires lower per-iteration computational complexity.

## 6. CONCLUSION

This paper seeks an improved version of the conventional WMMSE and FP without requiring the complexity of tuning the Lagrange multiplier. The proposed improvement is based on an interpretation of the conventional methods as the MM algorithm. We propose to further introduce a surrogate function bounding to make $\mathbf{W}$ much easier to update in the presence of power constraint. Our numerical result shows that the proposed method FP+ yields much faster convergence than the conventional FP.

## 7. REFERENCES

[1] P. C. Weeraddana, M. Codreanu, M. Latva-aho, A. Ephremides, and C. Fischione, "Weighted sum-rate maximization in wireless networks: A review," *Found. Trends Netw.*, vol. 6, no. 1–2, pp. 1–163, 2012.

[2] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, 2008.

[3] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, 2011.

[4] K. Shen and W. Yu, "Fractional programming for communication systems—Part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, 2018.

[5] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, 2008.

[6] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1142–1157, 2010.

[7] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, 2016.

[8] X. Zhao, S. Lu, Q. Shi, and Z.-Q. Luo, "Rethinking WMMSE: Can its complexity scale linearly with the number of BS antennas?" *arXiv preprint arXiv:2205.06225*, 2022.

[9] I. M. Stancu-Minasian, *Fractional programming: Theory, methods and applications*. Springer Sci. & Bus. Media, 2012, vol. 409.

[10] K. T. K. Cheung, S. Yang, and L. Hanzo, "Achieving maximum energy-efficiency in multi-relay OFDMA cellular networks: A fractional programming approach," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2746–2757, 2013.

[11] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory," *Found. Trends Commun. Inf. Theory*, vol. 11, no. 3-4, pp. 185–396, 2015.

[12] D. P. Bertsekas, *Nonlinear programming*. Belmont, MA, USA: Athena Scientific, 1999.

[13] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, 2020.

[14] A. A. Khan, R. S. Adve, and W. Yu, "Optimizing downlink resource allocation in multiuser MIMO networks via fractional programming and the Hungarian algorithm," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5162–5175, 2020.

[15] K. Shen, H. V. Cheng, X. Chen, Y. C. Eldar, and W. Yu, "Enhanced channel estimation in massive MIMO via coordinated pilot design," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 6872–6885, 2020.

[16] S.-H. Park, S. Jeong, J. Na, O. Simeone, and S. Shamai, "Collaborative cloud and edge mobile computing in C-RAN systems with minimal end-to-end latency," *IEEE Trans. Signal and Inf. Process. over Netw.*, vol. 7, pp. 259–274, 2021.

[17] K. Shen and W. Yu, "Fractional programming for communication systems—Part II: Uplink scheduling via matching," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2631–2644, 2018.

[18] K. Shen, W. Yu, L. Zhao, and D. P. Palomar, "Optimization of MIMO device-to-device networks via matrix fractional programming: A minorization–maximization approach," *IEEE/ACM Trans. Netw.*, vol. 27, no. 5, pp. 2164–2177, 2019.

[19] Z. Zhang and Z. Zhao, "Weighted sum-rate maximization for multi-hop RIS-aided multi-user communications: A minorization-maximization approach," in *22nd Int. Workshop on Signal Process. Adv. in Wireless Commun. (SPAWC)*, 2021, pp. 106–110.

[20] ——, "Rate maximizations for reconfigurable intelligent surface-aided wireless networks: A unified framework via block minorization-maximization," *arXiv preprint arXiv:2105.02395*, 2021.

[21] L. Pellaco, M. Bengtsson, and J. Jaldén, "Matrix-inverse-free deep unfolding of the weighted MMSE beamforming algorithm," *IEEE Open J. Commun. Society*, vol. 3, pp. 65–81, 2021.