Frontiers of Science Awards
for TCIS
pp. 1–12

A Perspective on Fractional Programming for Communications System Design

Kaiming Shen and Wei Yu

Abstract

Fractional programming (FP) refers to a class of optimization problems with ratio term(s) as a main component. This article provides a perspective on some recent advances in the algorithm development for FP, as well as the application of FP to power control for communications system design. The main focus of this article is a recently developed technique called quadratic transform for tackling multiple-ratio FP problems—in contrast to classic FP techniques that are typically limited to the single-ratio case. Multiple-ratio problems are important for the optimization of communication networks, because system-level design in communications and networking often involves multiple signal-to-interference-plus-noise ratios (SINRs). Moreover, we survey some recent theoretical advances in quadratic transform.

1 Introduction

Fractional programming (FP) refers to a class of optimization problems containing ratio terms. Its history can be traced back to an early paper on economic expansion [Neu] by John von Neumann in the 1930's; it has since inspired extensive study in broad areas of economics, management science, information theory, optics, graph theory, and computer science [Sta, Baj].

The aim of this article is to provide a perspective on the recent development of a new technique for solving FP problems in [SY1] and its applications to communication system design, in particular to the important problem of power control for wireless systems. Although an extensive prior literature already exists for FP, most of them are applicable only to single-ratio problems. For example, early works on FP for communications system design [ZaJ, IC$^+$], which rely on the classical techniques of Charnes-Cooper algorithm [ChC, Sch] and Dinkelbach's algorithm [Din], have had to limit the system models to scenarios involving only one single ratio. However, modern system-level communication network design often involves multiple ratios, because the overall system performance is typically a function of multiple fractions—such as the signal-to-interference-plus-noise ratios (SINRs) of multiple interfering links. Yet, because the multiple-ratio FP problem is NP-hard [Cro], the prior works [Ben, Kun, PhT] on multiple-ratio FP mostly resort to the branch-and-bound algorithms, which have exponential complexity.

This article considers the multiple-ratio FP from a new perspective using techniques first proposed in [SY1, SY2]. The main theoretic contribution of [SY1] is a novel technique called the quadratic transform that introduces a set of suitable auxiliary variables, then recasts the original problem to a form amenable to iterative optimization. Specifically, this new technique decouples the numerator and the denominator of each ratio term, similar to the classical Dinkelbach's transform [Din], but works with multiple ratios as opposed to (almost exclusively) single-ratio problems for the classic method. This decoupling feature of the newly proposed quadratic transform is particularly suited for coordinating the SINRs across multiple interfering links in wireless networks.

2 FP Problems

2.1 Single-Ratio FP

We start with the simplest case of FP, the single-ratio problem. Consider a pair of nonnegative function $A(x) \geq 0$ and strictly positive function $B(x) > 0$, where the variable x is restricted to a nonempty constraint set $\mathcal{X}$. Assume also that $A(x)$ and $B(x)$ are both differentiable. The *single-ratio* problem is defined to be

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \frac{A(x)}{B(x)} \\ \text{subject to} \quad & x \in \mathcal{X}. \end{aligned} \tag{2.1}$$

In the FP literature, $A(x)$ is often assumed to be a concave function, while $B(x)$ is assumed to be convex; further, $\mathcal{X}$ is assumed to be a convex set. This set of additional assumptions are called the *concave-convex* condition. It turns out that the concave-convex condition is satisfied in many applications of FP. Note that the objective in (2.1) remains nonconcave, even under the concave-convex condition.

The difficulty in numerically solving the problem (2.1) is mainly due to the coupling between the numerator $A(x)$ and the denominator $B(x)$; so a natural idea is to decouple the ratio. The classic Dinkelbach's algorithm [Din] accomplishes this by rewriting the problem (2.1) as

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & A(x) - yB(x) \\ \text{subject to} \quad & x \in \mathcal{X} \end{aligned} \tag{2.2}$$

with an auxiliary variable $y \in \mathbb{R}$, which is iteratively updated as

$$y^{(t+1)} = \frac{A(x^{(t)})}{B(x^{(t)})}, \tag{2.3}$$

where t is the iteration index. For fixed y, the objective function of the new problem (2.2) is concave in x under the concave-convex condition, hence (2.2) can be efficiently solved by standard methods. As shown in [Din], the alternating steps between the convex optimization of x in (2.2) and the update of y in (2.3) guarantees convergence to a global optimum of the original problem (2.1).

Another classic method called the Charnes-Cooper algorithm [ChC, Sch] also seeks to decouple the ratio while convexifying the problem (2.1), albeit in a completely different way.

2.2 Multiple-Ratio FP

A natural extension of the single-ratio problem is the *sum-of-ratios* problem:

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \sum_{i=1}^{n} \frac{A_i(x)}{B_i(x)} \\ \text{subject to} \quad & x \in \mathcal{X}, \end{aligned} \tag{2.4}$$

where each $A_i(x) \geq 0$ and each $B_i(x) > 0$. The concave-convex condition can be extended accordingly: each $A_i(x)$ is concave, each $B_i(x)$ is convex, and $\mathcal{X}$ is a convex set. Unfortunately, neither the Dinkelbach's algorithm [Din] nor the Charnes-Cooper algorithm [ChC, Sch] can be extended for the sum-of-ratios problem (2.4). It is tempting to decouple each ratio $A_i(x)/B_i(x)$ individually as in (2.2) by using an auxiliary variable y_i, but the resulting new problem is not equivalent to the original problem. This is to say:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \sum_{i=1}^{n} \frac{A_i(x)}{B_i(x)} \quad \not\Leftrightarrow \quad \underset{x\in\mathcal{X}}{\text{maximize}} \sum_{i=1}^{n} \Big(A_i(x) - y_i B_i(x)\Big),$$

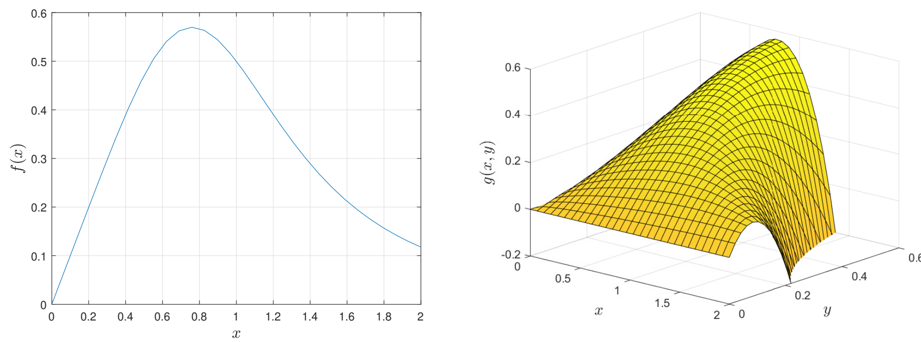
where each y_i is iteratively updated as $A_i(x)/B_i(x)$.

The reason why Dinkelbach's algorithm fails to work for problems involving multiple ratios is that it is not an embedding of $A(x)/B(x)$ into a higher dimensional function $g(x, y)$. To explain this more precisely, let us temporarily restrict attention to the single-ratio problem (2.1). Ideally, we would like to have a ratio-decoupling transformation $g(x, y)$ that satisfies the following:

- C1: *(Decoupling)* The new objective has the form $g(x, y) = Z_1(A(x))Q_1(y) + Z_2(B(x))Q_2(y)$, where $Z_1(\cdot), Z_2(\cdot)$ and $Q_1(\cdot), Q_2(\cdot)$ are some scalar functions, and y is an auxiliary variable.
- C2: *(Same Solution)* Let $y^\star = \text{argmax}_y g(x, y)$ for each x. An optimal $x^\star$ that maximizes $A(x)/B(x)$ along with its corresponding $y^\star$ should also be a maximizer of $g(x, y)$.
- C3: *(Same Objective Value)* For any $x \in \mathcal{X}$ along with the corresponding maximizer $y^\star$, we have $g(x, y^\star) = A(x)/B(x)$.

All of the above goals are well-motivated. C1 stems from the classic Dinkelbach's algorithm that decouples $A(x)$ and $B(x)$ through y; C2 ensures that the new problem is equivalent to the original problem in terms of the optimal optimization variable after ratio decoupling; C3 imposes a stronger equivalence between the new problem and the original problem, as motivated by multiple-ratio problems where multiple objective values are added together.

The issue is that the Dinkelbach's algorithm [Din] is not an embedding that satisfies C2–C3. Specifically, the choice of $y^\star = A(x)/B(x)$ in Dinkelbach's algorithm gives $g(x, y^\star) = 0$. This is the fundamental reason that Dinkelbach's algorithm cannot be applied to the sum-of-ratios problem.



(a) original nonconcave objective $A(x)/B(x)$ (b) new objective $g(x, y)$

Figure 1: Let the optimizing variable $x \in \mathbb{R}^d$. The original FP objective $A(x)/B(x) : \mathbb{R}^d \to \mathbb{R}$ is nonconcave. The quadratic transform in (2.5) embeds the original objective in a higher dimension as $g : \mathbb{R}^{d+1} \to \mathbb{R}$, which is concave in x for each y, and concave in y for each x. An example of $x/(x^4 + 1)$ is shown.

In contrast, the quadratic transform as developed in [SY1] satisfies all of C1–C3. Furthermore, we envision an algorithm for iterative updating of x and y. To facilitate the maximization over y for each fixed x, we further impose the following condition that would allow y to be efficiently updated via convex optimization:

C4: *(Concavity)* The new objective $g(x, y)$ is a smooth concave function of y for fixed x, i.e., $\partial^2 g/\partial y^2 \leq 0$,

These four conditions C1–C4 give rise to the quadratic transform.

Theorem 2.1 (Quadratic Transform) *The quadratic transform*

$$g(x, y) = 2y\sqrt{A(x)} - y^2 B(x) \tag{2.5}$$

satisfies C1–C4. Further, if C4 is strengthened to require $\partial^2 g/\partial y^2$ to be independent of y, then any $g(x, y)$ satisfying C1–C4 must be of the form

$$g(x, y) = 2(t_1 y + t_2)\sqrt{A(x)} - (t_1 y + t_2)^2 B(x), \tag{2.6}$$

for some $t_1 \neq 0$ and $t_2 \in \mathbb{R}$. Thus, the quadratic transform (2.5) is the only possible such transformation, up to an affine transformation in y.

From a geometric perspective, the quadratic transform embeds the original objective function $A(x)/B(x)$ into a higher dimensional space by introducing an auxiliary variable y. The new objective $g(x, y)$ is now concave in each of x and y separately, and thus is easier to optimize, as illustrated in Figure 1. This is reminiscent of the idea of lift and project in semidefinite programming relaxation.

We can now apply the quadratic transform to the sum-of-ratios problem and recast (2.4) as

$$\begin{aligned} \underset{x,\,\boldsymbol{y}}{\text{maximize}} \quad & \sum_{i=1}^{n}\left(2y_i\sqrt{A_i(x)} - y_i^2 B_i(x)\right) \\ \text{subject to} \quad & x \in \mathcal{X}, \quad y_i \in \mathbb{R}, \quad i = 1, \ldots, n, \end{aligned} \tag{2.7}$$

where an auxiliary variable y_i is introduced for each ratio $A_i(x)/B_i(x)$, and $\boldsymbol{y}$ denotes the vector of auxiliary variables $(y_1, \ldots, y_n)$.

The above result can be generalized further. Consider a sequence of non-decreasing differentiable concave functions $f_i : \mathbb{R} \to \mathbb{R}$, for $i = 1, \ldots, n$. The *sum-of-functions-of-ratios* problem is the following:

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \sum_{i=1}^{n} f_i\left(\frac{A_i(x)}{B_i(x)}\right) \\ \text{subject to} \quad & x \in \mathcal{X}. \end{aligned} \tag{2.8}$$

Clearly, problem (2.4) is a special case of the problem (2.8). By the quadratic transform, (2.8) can be recast as

$$\begin{aligned} \underset{x,\,\boldsymbol{y}}{\text{maximize}} \quad & \sum_{i=1}^{n} f_i\left(2y_i\sqrt{A_i(x)} - y_i^2 B_i(x)\right) \\ \text{subject to} \quad & x \in \mathcal{X}, \quad y_i \in \mathbb{R}, \quad i = 1, \ldots, n. \end{aligned} \tag{2.9}$$

Algorithmically, the quadratic transform allows the optimization of x and $\boldsymbol{y}$ in an alternating fashion. When x is held fixed, each y_i can be determined as

$$y_i = \frac{\sqrt{A_i(x)}}{B_i(x)}. \tag{2.10}$$

When $\boldsymbol{y}$ is held fixed, the new objective function in (2.9) is concave in x under the concave-convex condition. Thus, the original nonconvex problem is turned into a *sequence* of convex optimizations. The following theorem analyzes the convergence.

Theorem 2.2 (Convergence Analysis) *For the sum-of-functions-of-ratios problem* (2.8) *under the concave-convex condition, the alternating optimization between x and $\boldsymbol{y}$ in the new problem* (2.9) *guarantees convergence to a stationary point of the problem* (2.8). *Moreover, the primal objective value $\sum_{i=1}^{n} f_i\big(\frac{A_i(x)}{B_i(x)}\big)$ is nondecreasing after each iteration.*

We remark that the convergence condition stated in the above theorem is milder than that of the block coordinate descent (BCD) method [Ber]. Observe that the alternating optimization between x and $\boldsymbol{y}$ in (2.9) can be thought of as BCD. As shown in [Ber], the BCD method guarantees convergence to a stationary point, but it requires the subproblem for each iterate to have a unique solution. This is true, e.g., when each $A_i(x)$ is strictly concave and each $B_i(x)$ is strictly convex. In contrast, the above theorem asserts that the concave-convex condition alone is already sufficient without requiring strict convexity. The reason is that the quadratic transform has a minorization-maximization (MM) interpretation, which has a weaker convergence condition as discussed in [SY+].

3 Application to Wireless Power Control

This section shows how the quadratic transform can be useful in communications system design. Specifically, we focus on the wireless power control problem, which contains multiple SINRs of interfering links. We illustrate how the quadratic transform can be performed in different ways and compare their performances.

3.1 Problem Statement

Consider n wireless links that use the same spectrum band. Use i or $j = 1, \dots, n$ to index the link. Denote by $G_{ii} \in (0, 1)$ the direct channel attenuation of link i, $G_{ij} \in (0, 1)$ the interference channel attenuation from link j to link i for $j \neq i$, p_i the transmit power of link i, and σ^2 the background Gaussian noise power. The SINR of link i, denoted as Γ_i, is given by

$$\Gamma_i = \frac{G_{ii} p_i}{\sigma^2 + \sum_{j=1, j\neq i}^{n} G_{ij} p_j}. \tag{3.1}$$

With interference treated as noise, we can evaluate the data rate of each link i by Shannon's capacity formula $\log(1 + \Gamma_i)$. Moreover, a nonnegative weight $\omega_i \geq 0$ is assigned to each link i in accordance with its priority. Assuming that the parameters $\{G_{ii}, G_{ij}, \sigma^2, \omega_i\}$ are fixed and known, the power control problem seeks the optimal power vector $\boldsymbol{p} = (p_1, \dots, p_n)$ that maximizes a weighted sum rate:

$$\begin{aligned} \underset{\boldsymbol{p}}{\text{maximize}} \quad & \sum_{i=1}^{n} \omega_i \log(1 + \Gamma_i) \\ \text{subject to} \quad & 0 \leq p_i \leq P, \quad i = 1, \dots, n, \end{aligned} \tag{3.2}$$

where P is a given power constraint. The power control problem (3.2) is a well-known difficult problem, due to the nonconcavity of the objective function.

3.2 Direct Quadratic Transform

Observing that each SINR term Γ_i satisfies the concave-convex condition and also that Γ_i is nested in a nondecreasing function $f_i(\Gamma_i) = \omega_i \log(1+\Gamma_i)$, we can readily perform the quadratic transform (2.9) to recast (3.2) as

$$\begin{aligned} \underset{\boldsymbol{p}, \boldsymbol{y}}{\text{maximize}} \quad & \sum_{i=1}^{n} \omega_i \log\left(1 + 2y_i\sqrt{G_{ii}p_i} - y_i^2\left(\sigma^2 + \sum_{j=1, j\neq i}^{n} G_{ij}p_j\right)\right) \\ \text{subject to} \quad & 0 \leq p_i \leq P, \quad y_i \in \mathbb{R}, \quad i = 1, \dots, n. \end{aligned} \tag{3.3}$$

We propose to optimize $\boldsymbol{p}$ and $\boldsymbol{y}$ alternately in the new problem. For fixed $\boldsymbol{p}$, each y_i is optimally determined as

$$y_i^\star = \frac{\sqrt{G_{ii} p_i}}{\sigma^2 + \sum_{j=1, j\neq i}^{n} G_{ij} p_j}. \tag{3.4}$$

For fixed $\boldsymbol{y}$, optimizing $\boldsymbol{p}$ in (3.3) is a convex problem. The above optimization method is summarized in Algorithm 1. By Theorem 2.2, Algorithm 1 converges to a stationary point of the power control problem (3.2).

Algorithm 1 Direct Quadratic Transform for Power Control

repeat
  Solve for $\boldsymbol{p}$ in the new problem (3.3) by the standard method
  Update $\boldsymbol{y}$ according to (3.4)
until the convergence criterion is satisfied

Algorithm 1 requires numerically solving a convex subproblem in each iteration. It would be desirable to optimize $\boldsymbol{p}$ in closed form. Toward this end, we propose another way to apply the quadratic transform in the next section.

3.3 Closed-Form Quadratic Transform

In [SY2], a Lagrangian dual transform is proposed to "move" the ratios to the outside of logarithms in (3.2), so as to rewrite (3.2) equivalently as

$$\begin{aligned}&\underset{\boldsymbol{p},\,\boldsymbol{\gamma}}{\text{maximize}} \quad \sum_{i=1}^{n}\omega_i\log(1+\gamma_i)-\sum_{i=1}^{n}\omega_i\gamma_i+\sum_{i=1}^{n}\frac{\omega_i(1+\gamma_i)G_{ii}p_i}{\sigma^2+\sum_{j=1}^{n}G_{ij}p_j}\\ &\text{subject to} \quad 0\le p_i\le P,\quad \gamma_i\ge 0,\quad i=1,\ldots,n,\end{aligned} \tag{3.5}$$

where an auxiliary variable γ_i is introduced for each ratio, and $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_n)$. For fixed $\boldsymbol{p}$, each γ_i is optimally determined as

$$\gamma_i^{\star}=\frac{G_{ii}p_i}{\sigma^2+\sum_{j=1,j\neq i}^{n}G_{ij}p_j}. \tag{3.6}$$

The optimization of $\boldsymbol{p}$ for fixed $\boldsymbol{\gamma}$ is an FP. As only the last term of the new objective contains $\boldsymbol{p}$, the optimization of $\boldsymbol{p}$ in (3.5) boils down to a sum-of-ratios problem. By the quadratic transform, problem (3.5) is further converted to

$$\begin{aligned}&\underset{\boldsymbol{p},\,\boldsymbol{\gamma},\,\boldsymbol{y}}{\text{maximize}} \quad \sum_{i=1}^{n}\omega_i\log(1+\gamma_i)-\sum_{i=1}^{n}\omega_i\gamma_i\\ &\qquad\qquad +\sum_{i=1}^{n}\left[2y_i\sqrt{\omega_i(1+\gamma_i)G_{ii}p_i}-y_i^2\left(\sigma^2+\sum_{j=1}^{n}G_{ij}p_j\right)\right]\\ &\text{subject to} \quad 0\le p_i\le P,\quad \gamma_i\ge 0,\quad y_i\in\mathbb{R},\quad i=1,\ldots,n.\end{aligned} \tag{3.7}$$

For fixed $\boldsymbol{p}$ and $\boldsymbol{\gamma}$, the optimal update of each y_i is

$$y_i^{\star}=\frac{\sqrt{\omega_i(1+\gamma_i)G_{ii}p_i}}{\sigma^2+\sum_{j=1}^{n}G_{ij}p_j}. \tag{3.8}$$

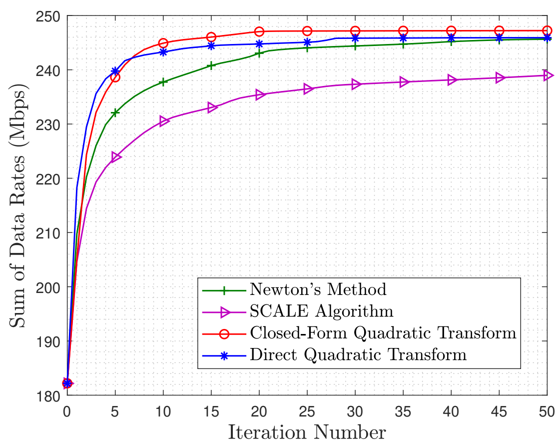


Figure 2: Maximizing the sum of downlink data rates across 7 cells in a wireless cellular network using different methods.

For fixed $\boldsymbol{\gamma}$ and $\boldsymbol{y}$, the optimal update of each p_i is

$$p_i^\star = \min\left\{P,\ \frac{\omega_i(1+\gamma_i)y_i^2G_{ii}}{\left(\sum_{j=1}^n y_j^2G_{ji}\right)^2}\right\}. \tag{3.9}$$

Thus all the variables are now optimized in closed form, as summarized in Algorithm 2.

Algorithm 2 Closed-Form Quadratic Transform for Power Control

1: **repeat**
2: update $\boldsymbol{\gamma}$ according to (3.6)
3: update $\boldsymbol{y}$ according to (3.8)
4: update $\boldsymbol{p}$ according to (3.9)
5: **until** the convergence criterion is satisfied

Figure 2 shows the convergence behaviors of the various power control algorithms in a 7-cell wrapped-around simulation environment with one user per cell. Here, we aim to maximize the sum of downlink data rates in a typical wireless cellular network. We use Newton's method and the SCALE algorithm [PaE] as benchmarks. Observe that the two FP algorithms have faster convergence as compared to the benchmarks.

4 Recent Advances

The past few years have witnessed considerable further theoretical developments in the quadratic transform along the following three directions:

- *Minimization FP:* The FP problem considered so far is limited to maximization of fractions. If we replace *max* by *min*, or place the ratios inside nonincreasing functions, the quadratic transform no longer works. Recently, an inverse quadratic transform [CZS] is developed to deal with minimization FP. Further, [CZS] proposes a unified quadratic transform for general FP problems wherein *max* and *min* coexist.
- *Matrix Ratio:* When the numerator function $A_i(x) \geq 0$ and the denominator function $B_i(x) > 0$ are generalized to be $m \times m$ matrix functions $\boldsymbol{A}_i(x) \succeq \mathbf{0}$ and $\boldsymbol{B}_i(x) \succ \mathbf{0}$, the scalar-valued ratio can be extended to a matrix ratio as

$$\frac{A_i(x)}{B_i(x)} \in \mathbb{R} \quad \Rightarrow \quad \boldsymbol{A}_i(x)\boldsymbol{B}_i^{-1}(x) \in \mathbb{R}^{m\times m}. \tag{4.1}$$

  Problem (2.4) can then be generalized to a sum-of-traces-of-matrix-ratios problem. A matrix-ratio extension of the quadratic transform is proposed in [SY$^+$]. However, it requires inverting an $m \times m$ matrix in each iteration, so its complexity can be high when m is large. The work [ZL$^+$] suggests ways to reduce the size of matrix under certain conditions; recently [ZZS] is able to eliminate the matrix inverse altogether to enable efficient implementation.
- *Convergence Rate Analysis and Acceleration:* From an MM theoretic viewpoint, [SZ$^+$] shows that the error bound in objective value diminishes at a rate of order $O(1/k)$ for the quadratic transform, where k is the number of iterates. Moreover, [SZ$^+$] establishes a connection between the quadratic transform and the gradient projection, based on which the heavy-ball method can be employed to accelerate the convergence. For this accelerated quadratic transform, its error bound can achieve a faster diminishing rate of $O(1/k^2)$.

5 Conclusion

The quadratic transform can tackle a broad range of FP problems with multiple ratios, whereas the classic methods are limited to the single-ratio FP only. Based on the quadratic transform, two different FP approaches are devised for solving the power control problem in wireless communications system design. These proposed methods recast the original nonconvex problem as a sequence of convex problems, thereby allowing efficient iterative optimization with provable convergence to a stationary point. Finally, various extensions of these ideas are discussed.

References

[Neu] J. von Neumann, *Uber ein okonomisches gleichgewichtssystem und eine verallgemeinerung des brouwerschen fixpunktsatzes*, Ergebnisse eines Mathematischen Kolloquiums **31** (1937), 73–83.

[Sta] I. M. Stancu-Minasian, *Fractional Programming: Theory, Methods and Applications*, Kluwer Academic Publishers (1992).

[Baj] E. B. Bajalinov, *Linear-Fractional Programming: Theory, Methods, Applications and Software*, Kluwer Academic Publishers (2003).

[SY1] K. Shen and W. Yu, *Fractional programming for communication systems—Part I: Power control and beamforming*, IEEE Trans. Signal Process. **66** (2018), 2616–2630.

[ZaJ] A. Zappone and E. Jorswieck, *Energy efficiency in wireless networks via fractional programming theory*, Foundations Trends Commun. Inf. Theory **11** (2015), 185–396.

[IC+] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, *Framework for link-level energy efficiency optimization with informed transmitter*, IEEE Trans. Wireless Commun. **11** (2012), 2946–2957.

[ChC] A. Charnes and W. W. Cooper, *Programming with linear fractional functionals*, Naval Research Logistics **9** (1962), 181–186.

[Sch] S. Schaible, *Parameter-free convex equivalent and dual programs of fractional programming problems*, Zeitschrift für Operations Research **18** (1974), 187–196.

[Din] W. Dinkelbach, *On nonlinear fractional programming*, Manage. Sci. **133** (1967), 492–498.

[Cro] J.-P. Crouzeix, *Algorithms for generalized fractional programming*, Mathematical Programming **52** (1991), 191–207.

[Ben] H. P. Benson, *Solving sum of ratios fractional programs via concave minimization*, J. Optimization Theory Appl. **135** (2007), 1–17.

[Kun] T. Kuno, *A branch-and-bound algorithm for maximizing the sum of several linear ratios*, J. Global Optimization **22** (2002), 155–174.

[PhT] N. T. H. Phuong and H. Tuy, *A unified monotonic approach to generalized linear fractional programming*, J. Global Optimization **22** (2003), 229–259.

[Ber] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific (2016).

[SY2] K. Shen and W. Yu, *Fractional programming for communication systems—Part II: Uplink scheduling via matching*, IEEE Trans. Signal Process. **66** (2018), 2631–2644.

[PaE] J. Papandriopoulos and J. S. Evans, *SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks*, IEEE Trans. Inf. Theory **55** (2009), 3711–3724.

[CZS] Y. Chen, L. Zhao, and K. Shen, *Mixed max-and-min fractional programming for wireless networks*, IEEE Trans. Signal Process. **72** (2023), 337–351.

[SY+] 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. **27** (2019), 2164–2177.

[SZ+] K. Shen, Z. Zhao, Y. Chen, Z. Zhang, and H. V. Cheng, *Accelerating* *quadratic transform and WMMSE*, IEEE J. Sel. Area Commun. (2024).

[ZL+] X. Zhao, S. Lu, Q. Shi, and Z.-Q. Luo, *Rethinking WMMSE: Can its complexity scale linearly with the number of BS antennas?* IEEE Trans. Signal Process. **71** (2023), 433–446.

[ZZS] Z. Zhang, Z. Zhao, and K. Shen, *Enhancing the efficiency of WMMSE and FP for beamforming by minorization-maximization*, Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (2023).

Kaiming Shen, School of Science and Engineering, The Chinese University of Hong Kong (Shenzhen), China.

E-mail address: shenkaiming@cuhk.edu.cn

Wei Yu, Electrical and Computer Engineering Department, University of Toronto, Canada.

E-mail address: weiyu@ece.utoronto.ca