Quadratic Transform for Fractional Programming in Signal Processing and Machine Learning

Kaiming Shen, *Senior Member, IEEE* and Wei Yu, *Fellow, IEEE*

***Abstract*—Fractional programming (FP) is a branch of mathematical optimization that deals with the optimization of ratios. It is an invaluable tool for signal processing and machine learning, because many key metrics in these fields are fractionally structured, e.g., the signal-to-interference-plus-noise ratio (SINR) in wireless communications, the Cramér-Rao bound (CRB) in radar sensing, the normalized cut in graph clustering, and the margin in support vector machine (SVM). This article provides a comprehensive review of both the theory and applications of a recently developed FP technique known as the *quadratic transform*, which can be applied to a wide variety of FP problems, including both the minimization and the maximization of the sum of functions of ratios as well as matrix-ratio problems.**

I. Introduction

FRACTIONAL programming (FP) refers to optimization problems involving ratios. Consider for example the maximization of the sum of multiple ratios over a vector $x \in \mathbb{R}^d$:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad f(x) := \sum_{i=1}^{n} \frac{A_i(x)}{B_i(x)}, \tag{1}$$

where $A_i(x) \geq 0$ and $B_i(x) > 0$ are functions of x. While one could simply treat $f(x)$ as a generic function and use, e.g., gradient-based method, to maximize the objective, better approaches are possible if we exploit the fractional structure of $f(x)$. Generic approaches for optimizing fractions do not always work well, because $B_i(x)$ can have strong influence on the overall objective value when $B_i(x)$ is close to zero. Consequently, it is not always easy to find the appropriate step size when optimizing the objective with the numerators $A_i(x)$ and the denominators $B_i(x)$ coupled in the fractional form. An important idea in the study of FP problems is to decouple the numerators and the denominators and to put them inside different terms in a summation. Specifically, as we wish to maximize $A_i(x)$ while minimizing $B_i(x)$ at the same time, it would be sensible to maximize the sum of an *increasing* function of $A_i(x)$ and a *decreasing* function of $B_i(x)$. The choices of these *transformed* objectives would affect the general applicability and performance of the methods. Many of the classical and modern FP techniques hinge on the judicious choices of these functions in the transform.

Manuscript submitted to IEEE Signal Processing Magazine on August 3, 2024, revised on December 21, 2024 and March 13, 2025, and accepted on March 24, 2025.

Kaiming Shen is with the School of Science and Engineering, The Chinese University of Hong Kong (Shenzhen), Shenzhen 518172, China (e-mail: shenkaiming@cuhk.edu.cn).

Wei Yu is with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: weiyu@ece.utoronto.ca).

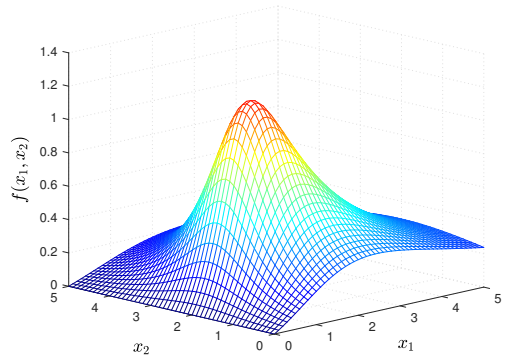


Fig. 1. An example of a single-ratio objective function $f(x_1,x_2) = \frac{x_1}{(x_1-1)^2+(x_2-1)^2+1}$ that satisfies the concave-convex condition and thus is quasi-concave, for which a local optimum is also the global optimum.

Classical methods in FP, such as the Charnes-Cooper method [1], [2], Dinkelbach's method [3], as well as the method in [4] that transforms the optimization of a rational function into a hierarchy of semidefinite programming (SDP) relaxations, all have such transforms embedded in their core ideas.

This article provides a comprehensive review of a recently developed technique called the quadratic transform [5] to solve FP problems. The main idea of quadratic transform is to introduce a set of auxiliary variables y_i and to transform the maximization of the sum of ratios over x as in (1) to a joint maximization over x and y_i as follows:

$$\underset{x\in\mathcal{X},\, y_1,\cdots,y_n}{\text{maximize}} \quad \sum_{i=1}^{n} \Big(2y_i\sqrt{A_i(x)} - y_i^2 B_i(x)\Big), \tag{2}$$

which can then be solved numerically by iterating between x and y_i's. The above equivalence can be established by explicitly optimizing over y_i for fixed x and substituting the optimal y_i back into the objective function. The key advantage of the quadratic transform is that it allows the decoupling of numerator and denominator for more than one ratio simultaneously in multi-ratio FP such as (1), whereas the classical FP methods such as the Charnes-Cooper and Dinkelbach's methods can decouple only a single ratio, as we shortly explain.

The study on FP was initiated by John von Neumann in 1937 in his seminal work on economic equilibrium. It has since been considered extensively in broad areas including economics, management science, information theory, optics, graph theory, and computer science [6]. Many early works focus on the single-ratio problem under the assumption that

TABLE I
DIFFERENT TYPES OF FP PROBLEMS AND THEIR APPLICATIONS

FP Problems	Quadratic Transform Methods	Examples of Applications
single-ratio: $\max_x A(x)/B(x)$	quadratic transform [5]	link-level energy efficiency
max-min-ratios: $\max_x \min_i A_i(x)/B_i(x)$	quadratic transform [5]	support vector machine (SVM), system-level energy efficiency
sum-of-ratios max: $\max_x \sum_i A_i(x)/B_i(x)$	quadratic transform [5]	power control in cellular networks
sum-of-ratios min: $\min_x \sum_i A_i(x)/B_i(x)$	inverse quadratic transform [10]	age-of-information (AoI) minimization
sum-of-functions-of-ratio: $\max_x \sum_i f_i(A_i(x)/B_i(x))$	unified quadratic transform [10]	secure transmission
sum-of-log-ratios: $\max_x \sum_i \log(1 + A_i(x)/B_i(x))$	logarithmic quadratic transform [11]	joint power control and user scheduling
matrix-ratio: extend $A_i(x)/B_i(x)$ to matrix $\sqrt{\boldsymbol{A}_i(x)}^{\mathrm{H}} \boldsymbol{B}_i^{-1}(x) \sqrt{\boldsymbol{A}_i(x)}$	matrix quadratic transform [12], nonhomogeneous quadratic transform [13], extrapolated quadratic transform [14]	unsupervised data clustering, multi-antenna beamforming, pilot design for channel estimation

the numerator is a concave function and the denominator is a convex function (known as the *concave-convex condition*), so that the optimization objective is quasi-concave, as shown by an example in Fig. 1. The quasi-concave structure of the single-ratio FP allows the development of classic methods, namely the Charnes-Cooper method [1], [2] and Dinkelbach's method [3] for achieving the global optimum of the single-ratio problem efficiently. These two classic methods have long been recognized as standard tools for FP. However, neither of them can be extended to the multi-ratio problems (except for the max-min-ratios case [7]). In fact, because even the basic sum-of-ratios problem can be shown to be NP-complete [8], many of the past studies for the multi-ratio FP have focused on the global optimization approaches such as branch-and-bound, which have an exponential worst-case complexity. Consequently, early applications of FP in communications and signal processing [9] are often restricted to either problems of small size, or problems that contain only a single ratio, e.g., the maximization of energy efficiency in a wireless network.

A basic tenet in optimization is that for an algorithm to be efficient, it should take advantage of the problem structure; generic optimization methods (e.g., gradient descent), which can be applied independently of the specific problem structure, would not be as efficient. By focusing attention on the fractional structure—which is a common characteristic of many communications and signal processing problems and is often the most important part of their main features, the quadratic transform is able to capture the problem-specific features while being generically applicable at the same time. Intuitively, the quadratic transform leverages the fractional structure of the problem while leaving other parts of the problem generic, in order to reach a desirable tradeoff between universality and exploitation of problem-specific structures.

This article starts by treating the single-ratio problem, the max-min-ratios problem, and the sum-of-ratios problem, then progresses toward a wider range of FP problems, including the sum-of-functions-of-ratio, the sum-of-log-ratios, and the matrix-ratio problems, as summarized in Table I along with their application areas. This feature article restricts attention to optimization problems with a single objective, although multi-objective FP [15] and bilevel FP [16] have also been considered in the literature. The focus of this article is on the quadratic transform and its relatives; we leave interested readers to consult additional literature [17], [18] for other heuristic approaches for solving FP, e.g., the harmony search and the evolutionary algorithm. Table II contains a summary of the features of the quadratic transform as compared to other commonly used optimization methods for FP problems with concave numerators and convex denominators.

In terms of the application area, the ubiquity of SINR is a main motivation for treating many communications and signal processing problems from an FP perspective. Unlike metrics such as the overall energy efficiency, SINRs are often measured at more than one entity (e.g., at multiple receivers or for different signals) in a communication system, so the corresponding problem is typically a multi-ratio FP. Recently, there have been a flurry of efforts in applying the quadratic transform to different research frontiers in wireless communications, e.g., cell-free massive multiple-input multiple-output (MIMO) system [19], satellite network [20], and intelligent reflecting surface (IRS) or reconfigurable intelligent surface (RIS) [21] systems. Other similar metrics, e.g., the signal-to-leakage-plus-noise ratio (SLNR) [22], can also be handled by the quadratic transform. Aside from the communications

TABLE II
COMPARISON OF THE DIFFERENT METHODS FOR FP WITH CONCAVE NUMERATORS AND CONVEX DENOMINATORS

	Charnes-Cooper [1], [2]	Dinkelbach [3], [7]	Quadratic Transform [5], [10]–[14]	AM-GM Inequality [31]
single-ratio	global optimum	global optimum	global optimum	global optimum
max-min-ratios	n/a	global optimum	global optimum	global optimum
sum-of-ratios max	n/a	n/a	stationary point	n/a
sum-of-ratios min	n/a	n/a	stationary point	stationary point
sum-of-log-ratios	n/a	n/a	stationary point	n/a
matrix-ratio	n/a	n/a	stationary point	n/a
convergence rate	iterations not required	superlinear	slower than Dinkelbach	slower than Dinkelbach

problems, the quadratic transform has been applied to the wireless sensing problems as well. In particular, there has been extensive recent research interest in using the quadratic transform to maximize the communication SINR and the radar SINR jointly for integrated sensing and communications (ISAC) applications, e.g., in [23]. Other commonly used metrics for radar signal processing include the CRB [10] and the ambiguity function sidelobe level ratio [24], both of which are fractionally structured and hence amenable to the quadratic transform based optimization. In the area of image processing, [25] uses Dinkelbach's method to handle the spectral level ratio for medical imaging, [26] uses Dinkelbach's method to maximize the ratio of data consistency to the regulation term for electrical capacitance tomography, and [27] proposes a novel FP approach to the regularized total least squares (RTLS) problem for image deblurring. Latency is another key metric with a fractional structure; the quadratic transform has been adopted to reduce latency in cloud radio access networks (C-RAN) [28] and in federated edge learning systems [29].

Machine learning is a field of ever increasing importance where FP has also found ample applications. SVM is a fundamental and popular supervised classifier. It aims to maximize the distance between the boundary and the nearest data points, named *margin*, in order to minimize the misclassification error. The margin maximization problem is nonconvex. A typical solution technique is to reformulate the problem into a convex form, then apply the Lagrangian duality theory. But since the margin has a fractional structure, the SVM problem can be directly solved by Dinkelbach's method. A potential advantage of using the FP approach for SVM is that it may be able to deal with multi-class SVM problems (which is a multi-ratio FP [30]), for which the classical method cannot be easily applied.

Aside from supervised classification, the unsupervised clustering problem is also closely related to FP, because the commonly used normalized-cut objective has a fractional structure. The authors of [32] focus on the two-class clustering and formulate the optimization objective as a single ratio. In contrast, [33] concerns the general multi-class clustering problem which has a sum-of-ratios optimization objective for which the quadratic transform can be used to decouple the multiple ratios to facilitate iterative optimization. Other recent applications of FP in machine learning include the submodular balanced clustering [34] and the fractional loss function approximation for federated learning [35].

It is worth pointing out that conversely the practical applications have also pushed the FP theory forward. Two specific examples follow. First, owing to the celebrated Shannon's capacity formula $C = \log(1+\mathrm{SINR})$, the work [11] proposes a novel FP technique termed the *Lagrangian dual transform* to address the log-ratio problem. Second, to account for MIMO transmission, a line of studies [10], [12], [36] develop a matrix generalization of the traditional scalar-valued FP to account for the ratio between two matrices. Moreover, extensive connections have been discovered between the FP method and other existing optimization methods for communications and signal processing, e.g., the fixed-point iteration, the weighted minimum mean squared error (WMMSE) algorithm, the minorization-maximization or majorization-minimization (MM) method, and the gradient projection method. In fact, the latest advances in the FP field mirror the newest frontiers in signal processing. For instance, as opposed to the conventional FP study that considers the ratio maximization and the ratio minimization separately, the recent work [10] aims at a unified approach to solve the mixed max-and-min FP problems, as motivated by the emerging application of ISAC, where the maximization of the SINRs and the minimization of the CRB coexist. As such, it is worthwhile to look at the state-of-the-art FP techniques in conjunction with their latest applications.

The rest of the article is organized according to the classification of the various FP problems. We begin with the single-ratio FP and the max-min-ratios FP, both of which can be solved by the classic methods. Next, we focus on the sum-of-ratios FP. Since the classic methods no longer work for the sum-of-ratios FP, a new method called the quadratic transform

is introduced. In particular, the maximization case and the minimization case of the sum-of-ratios FP need to be treated differently. As a generalization of the sum-of-ratios FP, we further discuss the sum-of-functions-of-ratio FP. We then pay special attention to the sum-of-logarithmic-ratios FP, because of the key role it plays in communication system design. Moreover, we discuss the matrix-ratio FP. Lastly, we connect the quadratic transform to other optimization methods and also analyze the rate of convergence.

Notation: We denote by $\mathbb{R}$ the set of real numbers, $\mathbb{C}$ the set of complex numbers, and $\mathbb{S}_{+}^{m\times m}$ (resp. $\mathbb{S}_{++}^{m\times m}$) the set of $m \times m$ positive semi-definite (resp. definite) matrices. We denote by $\|\cdot\|_2$ the Euclidean norm, and $\|\cdot\|_{\mathrm{F}}$ the Frobenius norm. For a matrix $\boldsymbol{A}$, let $\boldsymbol{A}^{\mathrm{H}}$ be its conjugate transpose and $\boldsymbol{A}^{\top}$ be its transpose. For a square matrix $\boldsymbol{A}$, let $\boldsymbol{A}^{-1}$ be its inverse (assuming that $\boldsymbol{A}$ is nonsingular) and $\mathrm{Tr}(\boldsymbol{A})$ be its trace. Denote by $\boldsymbol{I}$ the identity matrix. For a real number a, let $[a]_+ = \max(a, 0)$. Moreover, we use a letter without subscript to denote a set of variables over the subscripts, e.g., p as $\{p_i\}_{i=1}^n$, and $\mathbf{Y}$ as $\{\mathbf{Y}_i\}_{i=1}^n$.

II. Single-Ratio Problem

Consider a pair of numerator function $A(x)$ and denominator function $B(x)$. Assume that $A(x) \geq 0$ and $B(x) > 0$ with the variable x restricted to a nonempty set $\mathcal{X}$. We seek the optimal $x \in \mathcal{X}$ to maximize an objective that has a ratio form:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \frac{A(x)}{B(x)}. \tag{3}$$

Unless stated otherwise, we assume by convention that $A(x)$ is concave in x, $B(x)$ is convex in x, and $\mathcal{X}$ is a nonempty convex set, i.e., the *concave-convex condition* [5]. Note that problem (3) is still nonconvex in general under the concave-convex condition.

The coupling between $A(x)$ and $B(x)$ is the main difficulty in solving the problem (3). A natural idea is to decouple the ratio. This constitutes the fundamental basis of most FP methods (including the quadratic transform). For instance, the classic *Charnes-Cooper method* [1], [2] rewrites problem (3) as

$$\begin{aligned} \underset{z,\, q}{\text{maximize}} \quad & zA\left(\frac{q}{z}\right) \\ \text{subject to} \quad & zB\left(\frac{q}{z}\right) \leq 1 \\ & z \in \mathcal{Z},\ q \in \mathcal{Q}, \end{aligned} \tag{4}$$

with two new variables introduced as

$$z = \frac{1}{B(x)} \quad \text{and} \quad q = \frac{x}{B(x)}, \tag{5}$$

where the constraint sets $\mathcal{Z}$ and $\mathcal{Q}$ are determined over all $x \in \mathcal{X}$. Observe that only $A(x)$ is kept in the new objective function, while $B(x)$ is moved to the constraint. Most importantly, since $A(x)$ is concave and $B(x)$ is convex, the new problem is jointly convex in (z, q), so it can be efficiently solved by standard methods. After obtaining the optimal $(z^\star, q^\star)$, we can immediately recover the optimal solution of the original problem (3) as $x^\star = q^\star/z^\star$ according to (5).

In comparison, the ratio decoupling technique of the classic *Dinkelbach's method* [3] is more straightforward. It breaks down problem (3) into a sequence of subproblems

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad A(x) - yB(x), \tag{6}$$

where the auxiliary variable y is iteratively updated as

$$y = \frac{A(x)}{B(x)}. \tag{7}$$

Under the concave-convex condition, the subproblem (6) is convex in x and hence can be efficiently solved for fixed y. By solving a sequence of subproblems (6) with y iteratively updated as in (7), Dinkelbach's method guarantees that the sequence of solutions x converges to the optimal solution $x^\star$ in (3).

Thus, both the Charnes-Cooper method and Dinkelbach's method can attain the global optimum of problem (3) despite its nonconvexity. This is unsurprising, because the problem (3) is *quasi-convex* (and further *pseudo-convex* [6]) under the concave-convex condition. The following example shows an early application of the traditional single-ratio FP methods in wireless communications.

Example 1 (Energy Efficiency of Wireless Link): Consider a single wireless link. Denote by p the transmit power, $h \in \mathbb{C}$ the channel gain, and σ^2 the noise power. The channel capacity is computed as $\log(1 + |h|^2 p/\sigma^2)$. Aside from p, the wireless transmission system requires a constant ON-power of $\delta > 0$. We seek the optimal p that maximizes the energy efficiency [9], [37]:

$$\begin{aligned} \underset{p}{\text{maximize}} \quad & \frac{\log(1 + |h|^2 p/\sigma^2)}{p + \delta} \\ \text{subject to} \quad & 0 \leq p \leq P, \end{aligned} \tag{8}$$

where P is the max power constraint. Observe that the above single-ratio problem satisfies the concave-convex condition, so both the Charnes-Cooper method and Dinkelbach's method are applicable.

We remark that not all single-ratio problems in the literature meet the concave-convex condition. In the above example, when there are multiple wireless links (so each link has its own power variable), the numerator becomes the sum rate which is no longer a concave function of the power variables, so the new problem after applying Dinkelbach's method is still nonconvex. Further optimization methods are required after decoupling the ratio, as discussed in [9], [11]. As another example, the RTLS problem for image deblurring [27] aims to maximize the ratio between two convex quadratic functions. In this case, as shown in [9], decoupling the ratio (e.g., by Dinkelbach's method) is still beneficial even when the concave-convex condition does not hold; the FP technique can be applied in conjunction with other methods to facilitate optimization.

An important class of the single-ratio FP is the *rational optimization* [4], wherein both $A(x)$ and $B(x)$ are polynomial functions of $x \in \mathbb{R}^d$. The problem is closely related to the polynomial sum of squares and *Hilbert's 17th Problem* [4]. The state-of-the-art approach in this case is to consider an SDP approximation parameterized by an integer $r \geq 0$,

based on the epigraph lifting [4] or the generalized moment problem [38]. When r is increased, the SDP approximation becomes tighter and its solution eventually becomes exactly the solution of the rational optimization problem at a finite order [39] (although the complexity of SDP also becomes higher as r increases). This is known as *Lasserre's hierarchy* [40]. Furthermore, Lasserre's hierarchy can be extended for the sum-of-rational-functions optimization [38].

III. Max-Min-Ratios Problem

We now consider FP problems comprising more than one ratio, writing the numerator and denominator of the ith ratio term as $A_i(x)$ and $B_i(x)$, respectively. Similar to the single-ratio case, each numerator $A_i(x)$ is assumed to be nonnegative, while each denominator $B_i(x)$ is assumed to be strictly positive. Moreover, the concave-convex condition for the multi-ratio FP now means that each $A_i(x)$ is a concave function, each $B_i(x)$ is a convex function, and $\mathcal{X}$ is a nonempty convex set. We begin with the max-min-ratios problem:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \min_i \left\{ \frac{A_i(x)}{B_i(x)} \right\}. \tag{9}$$

Like the single-ratio case, the max-min-ratios FP problem is still nonconvex even under the concave-convex condition, so solving it directly is difficult.

As a classic result on the multi-ratio FP [7], Dinkelbach's method can be generalized for the max-min-ratios problem by rewriting (9) as

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \min_i \left\{ A_i(x) - yB_i(x) \right\}, \tag{10}$$

where the auxiliary variable y is iteratively updated as

$$y = \min_i \left\{ \frac{A_i(x)}{B_i(x)} \right\}. \tag{11}$$

It is then possible to reach the global optimum of the original problem (9) by solving the new problem (11) with y updated iteratively as above. We now present an application of this technique to a classic problem in machine learning.

Example 2 (Margin Maximization for SVM): SVM is an important tool for data classification and regression. The core of SVM lies in the minimization of the margin, which turns out to be a max-min-ratios problem. For ease of discussion, consider the binary classification of linearly separable data points. Given a set of data points $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$, where each $\boldsymbol{x}_i \in \mathbb{R}^m$ has a binary label $t_i \in \{+1, -1\}$, for $i = 1, 2, \ldots, n$, the aim of the classifier is to draw a decision boundary $y = \boldsymbol{w}^\top \boldsymbol{x} + b$ to divide the data points, i.e., label "+1" if $y > 0$ and "−1" otherwise. The distance from the data point $\boldsymbol{x}_i$ to the decision boundary, denoted by d_i, can be computed as

$$d_i = \frac{t_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b)}{\|\boldsymbol{w}\|_2}. \tag{12}$$

The shortest distance d_i across all the data points is called the *margin*, as shown in Fig. 2. Assuming that the data points are separable by a hyperplane, the objective of SVM is to seek the optimal decision boundary that maximizes the margin:

$$\underset{\boldsymbol{w}\in\mathbb{R}^m,\, b\in\mathbb{R}}{\text{maximize}} \quad \min_i \{d_i\}. \tag{13}$$

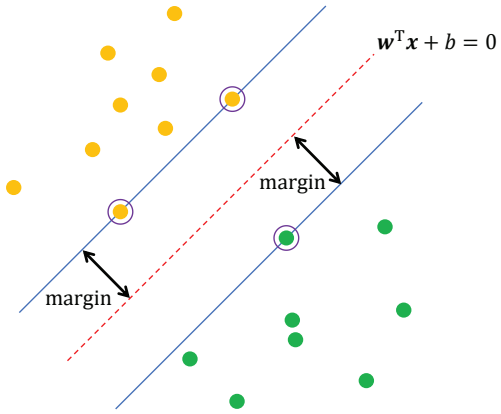


Fig. 2. The yellow points are labeled "−1"; the green points are labeled "+1". The red dashed line is the decision boundary. The margin maximization problem of SVM is to optimize the decision boundary to maximize the margin. Because the margin is a fractional function, the SVM optimization is a max-min-ratios problem.

The objective function here is a ratio, which is nonconvex. The usual treatment in most standard textbooks on this subject proceeds to transform the problem into an equivalent convex quadratic form. Subsequently, Lagrangian duality theory is applied to solve the problem in the dual domain.

Here, we point out that since each d_i is a fractional function, problem (9) can be directly treated as a max-min-ratios problem. In particular, (9) satisfies the concave-convex condition, so it can be immediately solved by the generalized Dinkelbach's method [7]. Aside from using this max-min-ratios FP approach, we can alternatively regard the entire $\min_i\{t_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b)\}$ as the numerator and notice that the resulting single-ratio problem meets the concave-convex condition, so the Charnes-Cooper method [1], [2] and Dinkelbach's method [3] are also applicable here.

We remark here that the max-min-ratios FP is also involved in the so-called Grab-n-Pull framework with broad applications for relay beamforming, Doppler-robust waveform design for active sensing, and robust data classification [41].

IV. Sum-of-Ratios Problem

As compared to the max-min-ratios problem, the sum-of-ratios problems are much more challenging. Further, we also need to deal with maximization and minimization separately, as elaborated below. It is known [8] that both the maximization and minimization of the sum of ratios are NP-complete, even when the numerators and denominators are all linear functions. For example, a sum-of-ratios problem with 20 ratios already cannot be globally solved within reasonable time according to [42]. As such, finding a local optimum efficiently is what one can expect at best.

The difference between maximization and minimization adds a further complication to the sum-of-ratios problems. The previous section only discusses how to maximize the objective function for the single-ratio problem (3) or the maximization of min-ratios (9). If instead we consider the minimization of a

single-ratio or the minimization of max-ratios, it would suffice to take the reciprocal(s) of the ratio term(s) and apply the same technique as before (with the concave-convex condition reversed as the convex-concave condition). However, when it comes to the sum-of-ratios problem, we can no longer convert the minimization to maximization by simply taking the reciprocals. For this reason, we discuss the sum-of-ratios maximization problem and the sum-of-ratios minimization problem separately.

A. Sum-of-Ratios Maximization Problem

We start with the sum-of-ratios maximization problem:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \sum_{i=1}^{n} \frac{A_i(x)}{B_i(x)}. \tag{14}$$

The natural extension of Dinkelbach's method from the single-ratio to the max-min-ratios problem as discussed earlier may lead one into believing that other multi-ratio problems can be dealt with in a similar fashion. Unfortunately, the max-min-ratios problem is a rare case, whereas most multi-ratio problems cannot be addressed by extending Dinkelbach's method. For instance, it is tempting to generalize Dinkelbach's method for the above sum-of-ratios problem by using auxiliary variables y_i updated as $A_i(x)/B_i(x)$ in an iterative fashion, then maximizing a transformed problem in each step as in the single-ratio Dinkelbach's method. But this is not equivalent to the original problem (14). In other words,

$$\underset{x\in\mathcal{X}}{\text{maximize}} \sum_{i=1}^{n} \frac{A_i(x)}{B_i(x)} \not\Leftrightarrow \underset{x\in\mathcal{X}}{\text{maximize}} \sum_{i=1}^{n} \Big(A_i(x) - y_i B_i(x)\Big), \tag{15}$$

when each y_i is iteratively updated as $y_i = A_i(x)/B_i(x)$. As explained in [5], the fundamental reason for the breakdown of Dinkelbach's method in the multi-ratio case is that the problem transformation lacks the *objective value equivalence*. This is to say that, although the problem (3) and the problem (6) have the same solution set, their objective values are not equal at the optimum. Thus, one cannot add multiple single-ratios together and expect the transformed problem to be equivalent to the original problem. For specific problems, e.g., energy efficiency problem, [43] suggests rewriting the sum-of-ratios problem as a parameterized polynomial optimization problem, while [44] proposes a water-filling-type algorithm.

To fix this issue in general, we propose a new ratio-decoupling transform that imposes the objective value equivalence, i.e., the objective value of the new problem must be equal to that of the original problem at the optimum. Moreover, we require the new objective function to be concave in the auxiliary variable for ease of iterative update. Under the above two assumptions, it is shown in [5] that the new transform can take a quadratic function form. Specifically, the problem (14) can be rewritten as

$$\underset{x,\, y}{\text{maximize}} \quad \sum_{i=1}^{n} \Big(2y_i\sqrt{A_i(x)} - y_i^2 B_i(x)\Big). \tag{16}$$

This new transform is termed *quadratic transform* as first proposed in [5]. We discuss why the quadratic form is preferable to other function forms in the later subsection titled *Connection with MM Method*. The new problem (16) is amenable to alternating optimization. When x is held fixed, each y_i is optimally determined as

$$y_i^\star = \frac{\sqrt{A_i(x)}}{B_i(x)}. \tag{17}$$

When y is held fixed, solving for x in (16) is a convex problem under the concave-convex condition. Later in the article, we see that there are also applications in which the concave-convex condition does not hold, yet x can still be efficiently optimized for fixed y. Under the concave-convex condition (and also assuming differentiability of $A_i(x)$ and $B_i(x)$), the iterative optimization over x and y is guaranteed to converge to a stationary point of the original optimization problem (14).

An intuitive comparison between Dinkelbach's method [3] and the quadratic transform [5] is as follows. To maximize a ratio A/B, we need to increase A and decrease B simultaneously. Thus, A can be treated as the utility and B the penalty. Dinkelbach's method, $A - yB$, then boils down to a utility-minus-penalty strategy, where the auxiliary variable y serves as the price for the penalty. The quadratic transform method, $2y\sqrt{A} - y^2B$, combines the utility and penalty differently. Because the quadratic transform places A inside a concave operation (i.e., the square root), it is less aggressive than Dinkelbach's method in boosting A. As a consequence, the quadratic transform converges more slowly than Dinkelbach's method when solving a single-ratio problem, but it enjoys an advantage in that it can be applied to multi-ratio FP, whereas Dinkelbach's method cannot.

B. Sum-of-Ratios Minimization Problem

We proceed to the case of minimization of the sum-of-ratios problem, written in a slightly different form as below:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad -\sum_{i=1}^{n} \frac{A_i(x)}{B_i(x)}. \tag{18}$$

We write the minimization problem in the maximization form with the optimization objective multiplied by -1, in order to simplify the notation in the later developments. As already mentioned before, instead of the concave-convex condition in the maximization case, a common condition adopted for the minimization problem is that each $A_i(x)$ is a convex function, each $B_i(x)$ is a concave function, and $\mathcal{X}$ is still a nonempty convex set, namely the *convex-concave condition*. A naive idea is to merge -1 into $A_i(x)$ or $B_i(x)$ and thereby apply the previous FP method for the maximization problem; this is however problematic since the resulting new FP problem violates the condition that each numerator is nonnegative while each denominator is positive. Two methods emerged in the recent literature to handle the sum-of-ratios minimization problem.

The first method [10] aims to extend the quadratic transform to what is called the *inverse quadratic transform*. It recasts the

original problem (18) as

$$\underset{x\in\mathcal{X},\, y}{\text{maximize}} \quad -\sum_{i=1}^{n} \frac{1}{\Big[2y_i\sqrt{B_i(x)} - y_i^2 A_i(x)\Big]_+}. \tag{19}$$

Note that the denominator is placed inside $[\,\cdot\,]_+$ in (19) in order to rule out negative value in the denominator. This operation is critical to the equivalence between the problems (18) and (19), for otherwise letting $y_i \uparrow 0$ while fixing x would make the objective in (19) go to infinity, thus making the maximization problem degenerate. For the new problem (19), we again optimize x and y in an alternating fashion, which is guaranteed to reach a stationary point of the original problem. When x is held fixed, each y_i can be optimally updated as

$$y_i^\star = \frac{\sqrt{B_i(x)}}{A_i(x)}. \tag{20}$$

For fixed y, optimizing x in (19) is a convex problem under the convex-concave condition.

Another method, as recently proposed in [31], is called the arithmetic-mean geometric-mean (AM-GM) inequality transform. It rewrites problem (18) as

$$\underset{x\in\mathcal{X},\, y}{\text{maximize}} \quad -\sum_{i=1}^{n} \left(y_i A_i^2(x) + \frac{1}{4y_i B_i^2(x)} \right). \tag{21}$$

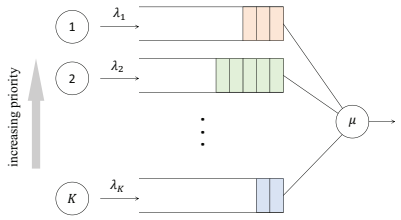
Like (19), the above new problem is amenable to the alternating optimization between x and y to reach a stationary point. Specifically, for fixed x, each y_i is optimally updated as

$$y_i^\star = \frac{1}{2A_i(x)B_i(x)}. \tag{22}$$

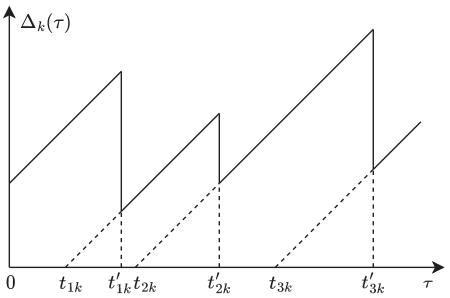
For fixed y, problem (21) is convex in x under the convex-concave condition, so it can be solved efficiently.

Although the inverse quadratic transform [10] and the AM-GM inequality transform [31] look quite different, it turns out that both of them can be placed under the umbrella of the MM theory [45], [46], so their convergence can be verified immediately. A later section is dedicated to this connection. We now present an application of the sum-of-ratios minimization FP.

Example 3 (Age-of-Information (AoI) Minimization in Queuing System): The notion of AoI [47] characterizes the freshness of data packet. Consider a K-sensor queuing network as depicted in Fig. 3(a). Each sensor k samples its source and delivers data packets at a rate λ_k (which is to be optimized), while the server processes the received packets at a constant rate μ, but treats the sensors with lower indices as higher priority, i.e., the packets from sensor k are served only if the queues associated with sensors $1, \cdots, k-1$ are all empty. The AoI for each sensor is defined as follows. The ith packet from the sensor k is delivered at time t_{ik} and departs the server at time t'_{ik}; the delay $t'_{ik} - t_{ik}$ is due to the waiting time in the queue and the server processing time. At the current time τ, let $\mathcal{N}_k(\tau)$ be the arrival time of the most recently received packet from the sensor k, i.e., $\mathcal{N}_k(\tau) = \max\{t'_{ik} : t'_{ik} \le \tau\}$. The instantaneous AoI of the source k at time τ is given by $\Delta_k(\tau) = \tau - \mathcal{N}_k(\tau)$. As a result, $\Delta_k(\tau)$ increases linearly with τ, and drops whenever a new packet departs the server,



(a) A K-source queuing system with priority.



(b) A typical instantaneous AoI Δ_k curve versus time.

Fig. 3. Rate control for minimizing AoI. The average AoI $\bar{\Delta}_k$ is the average area of the trapezoid below each tooth of the sawtooth curve. Since $\bar{\Delta}_k$ is a sum of ratios, the problem of optimizing the sensor data packet arrival rate to minimize the average AoI is a sum-of-ratios minimization problem.

so $\Delta_k(\tau)$ has a sawtooth profile along the time axis as shown in Fig. 3(b). We are interested in the average AoI in the long run $\mathbb{E}[\Delta_k] = \lim_{T\to\infty} \frac{1}{T}\int_0^T \Delta_k(\tau)d\tau$, which can be shown to be the average area of the trapezoid below each tooth of the sawtooth curve as in Fig. 3(b). For the M/M/1 queue model [47], the problem of minimizing the sum-of-AoI for the K sources can be formulated as

$$\begin{aligned} \underset{\lambda}{\text{minimize}} \quad & \sum_{k=1}^{K} \left(\frac{\hat{\rho}_k^2 + 3\hat{\rho}_k + 1}{\mu(1+\hat{\rho}_k)} + \frac{(\hat{\rho}_k+1)^2}{\mu\rho_k} \right) \\ \text{subject to} \quad & 0 \le \lambda_k \le \mu, \quad k = 1, \ldots, K, \end{aligned} \tag{23}$$

where $\rho_k = \lambda_k/\mu$ and $\hat{\rho}_k = \sum_{i=1}^{k-1}\rho_i$. This problem can be recognized as a sum-of-ratios minimization problem satisfying the convex-concave condition, so we can apply either the inverse quadratic transform [10] or the AM-GM inequality transform [31]. A numerical example is shown in Fig. 4, where the processing rate is fixed at $\mu = 1$ and the numbers of source nodes varies as $K = 3, 4, \ldots, 10$. Consider two benchmarks: (i) *Equal Rate Optimization* [48] that assumes all λ_k's are equal and then performs a one-dimensional search; (ii) *Max Rate Scheme* that sets each $\lambda_k = \mu$. Observe that the FP method achieves a much lower sum of AoIs and thus provides fresher information. Moreover, we remark that the AoI problem can be also tackled by the extended Lasserre's hierarchy [38], since the numerators and denominators are all polynomials.

V. Sum-of-Functions-of-Ratio Problem

We now consider a further extension of the sum-of-ratios optimization by investigating mixed functions of ratios. To this end, consider a sequence of nondecreasing functions f_i^+ :

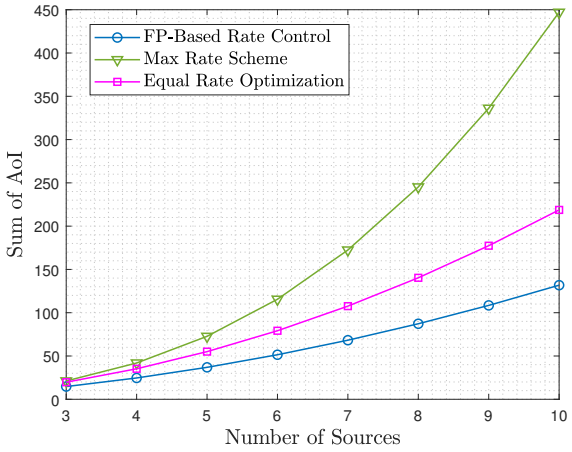


Fig. 4. Minimization of the sum of AoIs by different methods.

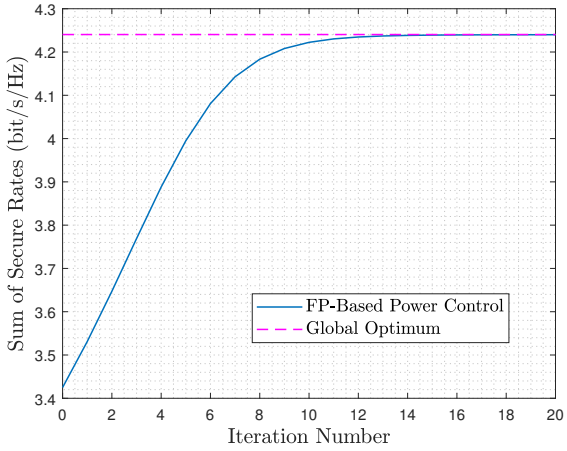


Fig. 5. Maximization of the sum of secure rates.

$\mathbb{R}_+ \to \mathbb{R}$ and a sequence of nonincreasing functions $f_i^- : \mathbb{R}_+ \to \mathbb{R}$, for $i = 1, 2, \ldots, n$, each of which takes a ratio term $A(x)/B(x)$ as input. We then arrive at a generalized sum-of-functions-of-ratio problem:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \sum_{i=1}^{n^+} f_i^+\left(\frac{A_i^+(x)}{B_i^+(x)}\right) + \sum_{i=1}^{n^-} f_i^-\left(\frac{A_i^-(x)}{B_i^-(x)}\right). \tag{24}$$

Intuitively, we aim to maximize the ratios inside $f_i^+(\cdot)$ and minimize the ratios inside $f_i^-(\cdot)$ at the same time. Thus, we can apply the quadratic transform to every ratio inside $f_i^+(\cdot)$ and in the meantime apply the inverse quadratic transform to every ratio inside $f_i^-(\cdot)$, thereby converting problem (24) to

$$\begin{aligned}\underset{x\in\mathcal{X},\, y,\, \tilde{y}}{\text{maximize}} \quad & \sum_{i=1}^{n^+} f_i^+\left(2y_i\sqrt{A_i^+(x)} - y_i^2 B_i^+(x)\right) \\ & - \sum_{i=1}^{n^-} f_i^-\left(\left[2\tilde{y}_i\sqrt{B_i^-(x)} - \tilde{y}_i^2 A_i^-(x)\right]_+^{-1}\right),\end{aligned} \tag{25}$$

where the auxiliary variables introduced by the quadratic transform are denoted by y_i, and the auxiliary variables introduced by the inverse quadratic transform are denoted by $\tilde{y}_i$. The preceding generalized quadratic transform is referred to as the *unified quadratic transform*.

As shown in a later part of the article, the unified quadratic transform is still an MM procedure, so the alternating optimization continues to guarantee convergence. Furthermore, the alternating optimization can be carried out efficiently provided that each ratio inside $f_i^+(\cdot)$ meets the concave-convex condition, while each ratio inside $f_i^-(\cdot)$ meets the convex-concave condition. When x is held fixed, all the auxiliary variables can be simultaneously optimally determined as

$$y_i^\star = \frac{\sqrt{A_i^+(x)}}{B_i^+(x)} \quad \text{and} \quad \tilde{y}_i^\star = \frac{\sqrt{B_i^-(x)}}{A_i^-(x)}. \tag{26}$$

When the auxiliary variables are held fixed, solving for x in (25) is a convex optimization problem. The following is an example of mixed max-and-min problem. It also shows that the choices of $f_i^+(\cdot)$ and $f_i^-(\cdot)$ are not necessarily unique for the same problem, and such choices may even be critical to the performance of the unified quadratic transform.

Example 4 (Secure Transmission): Consider the downlink of a wireless network with L cells. Within each cell, the base station (BS) transmits data to one legitimate user, at the risk of being wiretapped by an eavesdropper. Use $i = 1, 2, \ldots, L$ to index each cell and its associated legitimate user and eavesdropper. Denote by p_i the transmit power of BS i, $h_{ji} \in \mathbb{C}$ the channel from BS i to legitimate receiver j, $\tilde{h}_{ki} \in \mathbb{C}$ the channel from BS i to eavesdropper j, σ_i^2 the noise power at legitimate receiver i, and $\tilde{\sigma}_i^2$ the noise power at eavesdropper i. The SINRs of legitimate receiver i and eavesdropper i are respectively given by

$$\Gamma_i = \frac{|h_{ii}|^2 p_i}{\sum_{j\neq i}|h_{ij}|^2 p_j + \sigma_i^2} \quad \text{and} \quad \widetilde{\Gamma}_i = \frac{|\tilde{h}_{ii}|^2 p_i}{\sum_{j\neq i}|\tilde{h}_{ij}|^2 p_j + \tilde{\sigma}_i^2}. \tag{27}$$

We seek the optimal power allocation p that maximizes the total secrete data rates:

$$\begin{aligned}\underset{p}{\text{maximize}} \quad & \sum_{i=1}^{L}\Big(\log\left(1+\Gamma_i\right) - \log\left(1+\widetilde{\Gamma}_i\right)\Big) \\ \text{subject to} \quad & 0 \le p_i \le P, \quad i = 1, \ldots, L,\end{aligned} \tag{28}$$

where P is the power constraint on each BS. At first glance, the unified quadratic transform seems to be immediately applicable for problem (28) by letting $f_i^+(r) = \log(1+r)$ and $f_i^-(r) = -\log(1+r)$. But such $f_i^-(r)$ is not a concave function, so it is difficult to optimize p in the new problem when the auxiliary variables are fixed. The above issue can be resolved by rewriting the secrete data rate as

$$\begin{aligned}\underset{p}{\text{maximize}} \quad & \sum_{i=1}^{L}\Big(\log\left(1+\Gamma_i\right) + \log\Big(1 - \big(1+\widetilde{\Gamma}_i^{-1}\big)^{-1}\Big)\Big) \\ \text{subject to} \quad & 0 \le p_i \le P, \quad i = 1, \ldots, L,\end{aligned} \tag{29}$$

with $f_i^+(r) = \log(1+r)$ and $f_i^-(r) = \log(1-r)$. Observe that the new problem is convex in p for fixed $\{y_i, \tilde{y}_i\}$.

Fig. 5 shows some simulation results for this example. In

this simulation, we consider $L = 2$ cells. Let $\sigma_i^2 = -10$ dBm, $\tilde{\sigma}_i^2 = 0$ dBm, $P = 10$ dBm, $|h_{11}|^2 = 1$, $|h_{12}|^2 = 0.1$, $|h_{21}|^2 = 0.09$, $|h_{22}|^2 = 0.87$, $|\tilde{h}_{11}|^2 = 0.5$, $|\tilde{h}_{12}|^2 = 0.11$, $|\tilde{h}_{21}|^2 = 0.13$, and $|\tilde{h}_{22}|^2 = 0.39$. The global optimum is obtained via exhaustive search. Observe from Fig. 5 that the FP-based power control algorithm converges to the global optimum solution in this example. Observe also that the FP method has fast convergence.

VI. Sum-of-Logarithmic-Ratios Problem

Because of Shannon's capacity formula for the Gaussian channel, the following log-ratio problem deserves special attention:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \sum_{i=1}^{n} w_i \log\left(1+\frac{A_i(x)}{B_i(x)}\right), \tag{30}$$

where each ratio $A_i(x)/B_i(x)$ can be interpreted as the SINR of link i, and each nonnegative weight $w_i \geq 0$ can be interpreted as the priority for each link.

The above problem aims to maximize a sum of weighted rates across multiple links. One can immediately recognize problem (30) as a special case of problem (24), so the unified quadratic transform is applicable here. As a result, we can solve a sequence of convex problems in x with the auxiliary variables iteratively updated.

However, the above method has two downsides. First, although the unified quadratic transform can convert the log-ratio problem (30) into a sequence of convex subproblems in x, solving for x in each subproblem can only be done as a generic convex optimization problem as shown in the previous example in secure transmission. It would have been more desirable if it is possible to exploit the structure of these subproblems. Second, if we seek further extension of the FP technique to the case where the constraint set $\mathcal{X}$ is discrete, then decoupling ratios inside the logarithms does not help much, because we still face a challenging nonlinear discrete optimization problem after the transform.

Since the nonlinearity of logarithm is the reason for the above issues, a natural idea is to try to "move" ratios to the outside of the logarithms. Toward this end, the following *Lagrangian dual transform* has been developed in [11], where it is shown that (30) is equivalent to

$$\underset{x\in\mathcal{X},\,\gamma}{\text{maximize}} \quad \sum_{i=1}^{n} w_i \left(\log\left(1+\gamma_i\right)+\frac{(1+\gamma_i)A_i(x)}{A_i(x)+B_i(x)}-\gamma_i\right). \tag{31}$$

Here, an auxiliary variable γ_i is introduced for each ratio term $A_i(x)/B_i(x)$. Note that when moving $A_i(x)/B_i(x)$ to outside of the logarithm, we also need to add $A_i(x)$ to the denominator. For fixed x, each γ_i is optimally determined as

$$\gamma_i^\star = \frac{A_i(x)}{B_i(x)}. \tag{32}$$

For fixed γ, optimizing x in problem (31) boils down to a sum-of-weighted-ratios problem as formerly discussed. The next example shows that using the Lagrangian dual transform coupled with the quadratic transform can result in a sequence of subproblems with the sum-of-ratios structure, whereas using the quadratic transform alone cannot.

Example 5 (Power Control for Interfering Links): Consider K wireless links that reuse the same spectral band. Denote by $h_{ij} \in \mathbb{C}$ the channel from the transmitter of link j to the receiver of link i, p_i the transmit power of link i, P the power constraint, and σ^2 the background noise power. We consider the following power control problem of maximizing the sum of weighted rates:

$$\begin{aligned} \underset{p}{\text{maximize}} \quad & \sum_{i=1}^{K} w_i \log\left(1+\frac{|h_{ii}|^2 p_i}{\sum_{j\neq i}|h_{ij}|^2 p_j+\sigma^2}\right) \\ \text{subject to} \quad & 0 \leq p_i \leq P, \quad i=1,\ldots,K, \end{aligned} \tag{33}$$

where the weight $w_i > 0$ reflects the priority of link i. Although the unified quadratic transform enables a convex reformulation of the power control problem, updating p still requires solving a generic convex optimization problem.

We now show that using the Lagrangian dual transform coupled with the quadratic transform can lead to an FP problem in each iterate. First, by the Lagrangian dual transform, we move the ratios out of the logarithms:

$$\begin{aligned} \underset{p,\,\gamma}{\text{maximize}} \; & \sum_{i=1}^{K} w_i \left(\log(1+\gamma_i)+\frac{(1+\gamma_i)|h_{ii}|^2 p_i}{\sum_{j=1}^{K}|h_{ij}|^2 p_j+\sigma^2}-\gamma_i\right) \\ \text{subject to} \; & 0 \leq p_i \leq P, \quad i=1,\ldots,K. \end{aligned} \tag{34}$$

After γ_i's are optimally updated as in (32), solving for p in the above problem boils down to a sum-of-ratios maximization FP as considered previously in this article. As a result, each p_i can be optimally solved in closed form in each iteration as

$$p_i = \min\left\{P, \frac{y_i^2 w_i(1+\gamma_i)|h_{ii}|^2}{\left(\sum_{j=1}^{K} y_j^2 |h_{ji}|^2\right)^2}\right\}, \tag{35}$$

where each y_i is an auxiliary variable introduced by the quadratic transform and is iteratively updated as

$$y_i = \frac{\sqrt{w_i(1+\gamma_i)|h_{ii}|^2 p_i}}{\sum_{j=1}^{K}|h_{ij}|^2 p_j+\sigma^2}. \tag{36}$$

Fig. 6 validates the performance advantage of the proposed FP-based power control method in a 7-cell wrapped-around network. We aim to maximize the sum of downlink data rates. The BS-to-BS distance is 0.8 km, $P = 43$ dBm, $\sigma^2 = -100$ dBm, and the spectrum bandwidth is 10 MHz. The pathloss is modeled as $128.1 + 37.6\log_{10}(d) + \tau$ (in dB), where the distance d is in km, and the shadowing τ has a zero-mean Gaussian distribution with 8 dB standard deviation. To avoid the bias caused by the starting point, we try out the same set of random starting points for the different algorithms and average their performance. We use Newton's method and the geometric programming (GP) based SCALE method [49] as benchmarks. Observe that the two FP methods achieve competitive sum rates as the benchmarks, but the computational complexity of FP is much lower, because the updates in FP are in closed form.

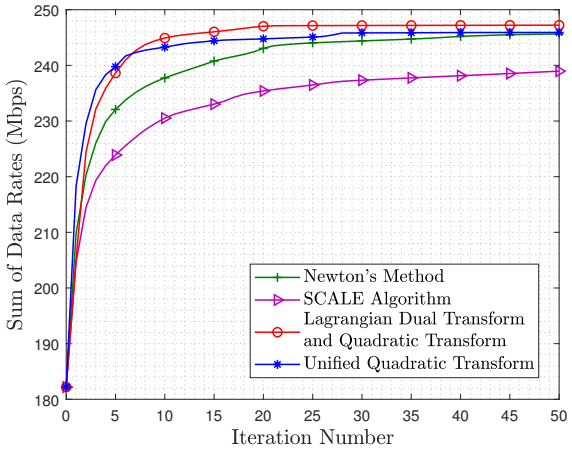


Fig. 6. Cumulative distributions of the sum rates in a 7-cell network achieved by the different power control methods.

VII. Matrix-Ratio Problems

A key advantage of the FP formulation is that it can be generalized to deal with matrix ratios. While the traditional setup of FP, as discussed thus far in the article, concerns a scalar ratio between a nonnegative function $A_i(x) \geq 0$ and a strictly positive function $B_i(x) > 0$, we now generalize it to a ratio between a positive semi-definite function $\boldsymbol{A}_i(x) \in \mathbb{S}_+^{m\times m}$ and a positive definite function $\boldsymbol{B}_i(x) \in \mathbb{S}_{++}^{m\times m}$. To lighten the notation, we drop the argument x in these matrix-valued functions in the rest of the article whenever doing so causes no confusion. Furthermore, we assume that each matrix numerator $\boldsymbol{A}_i$ can be factorized as $\boldsymbol{A}_i = \sqrt{\boldsymbol{A}_i}\sqrt{\boldsymbol{A}_i}^{\mathrm{H}}$ where $\sqrt{\boldsymbol{A}_i} \in \mathbb{C}^{m\times \ell}$, for some given positive integer ℓ. Note that the symbol $\sqrt{\boldsymbol{A}_i}$ denotes a matrix square root rather than an operation; note also that the above factorizations always exist so long as $\ell \geq \max_i\{\mathrm{rank}(\boldsymbol{A}_i)\}$. The traditional scalar-valued ratio is then generalized to the multidimensional case as below:

$$\frac{A_i}{B_i} \in \mathbb{R} \quad \Rightarrow \quad \sqrt{\boldsymbol{A}_i}^{\mathrm{H}} \boldsymbol{B}_i^{-1} \sqrt{\boldsymbol{A}_i} \in \mathbb{R}^{\ell\times\ell}. \tag{37}$$

One of the most intriguing aspects of the quadratic transform is that it can be extended to matrix ratios. Below we show how the extension works for the sum-of-ratios maximization case in (14).

To start, we write the matrix-ratio analog of the sum-of-ratios problem (14):

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \sum_{i=1}^{n} \mathrm{Tr}\Big(\sqrt{\boldsymbol{A}_i}^{\mathrm{H}} \boldsymbol{B}_i^{-1} \sqrt{\boldsymbol{A}_i}\Big). \tag{38}$$

The matrix ratios can be decoupled by a matrix extension of the quadratic transform:

$$\begin{aligned} \underset{x\in\mathcal{X},\, \boldsymbol{Y}}{\text{maximize}} \quad & \sum_{i=1}^{n} \mathrm{Tr}\Big(\sqrt{\boldsymbol{A}_i}^{\mathrm{H}} \boldsymbol{Y}_i + \boldsymbol{Y}_i^{\mathrm{H}} \sqrt{\boldsymbol{A}_i} - \boldsymbol{Y}_i^{\mathrm{H}} \boldsymbol{B}_i \boldsymbol{Y}_i\Big) \\ \text{subject to} \quad & \boldsymbol{Y}_i \in \mathbb{C}^{m\times\ell}, \quad i = 1, \ldots, n. \end{aligned} \tag{39}$$

The above matrix extension of the quadratic transform preserves the connection to the MM method, i.e., the alternating optimization between x and $\boldsymbol{Y}$ can still be recognized as an MM procedure, so the iterative process must converge. When x is held fixed, the auxiliary variables are optimally determined as

$$\boldsymbol{Y}_i^{\star} = \boldsymbol{B}_i^{-1} \sqrt{\boldsymbol{A}_i}. \tag{40}$$

For fixed $\boldsymbol{Y}$, solving for x in (39) is often much easier than the original problem (38) thanks to the matrix ratio decoupling. The following application of FP for data clustering illustrates this point. For the data clustering problem, neither the concave-convex condition nor the convex-concave condition is satisfied (because of the discrete constraints), but alternating optimization can still be performed efficiently after decoupling the ratio. Furthermore, we remark that the matrix extension can be considered for the more general sum-of-functions-of-ratio problem (24) as discussed in [10].

Example 6 (Normalized-Cut Problem for Data Clustering): Suppose that there are N data points in total. We use $i, j \in \{1, 2, \ldots, N\}$ as indices for these data points. For a pair of data points i and j, the similarity between them is quantified as $0 \leq w_{ij} \leq 1$. By symmetry, we have $w_{ij} = w_{ji}$. Visualizing in terms of a graph, the data points can be thought of as vertices, while the similarities can be thought of as edges with weight w_{ij} (or w_{ji}) assigned to the edge between vertex i and vertex j. Denote the set of vertices by $\mathcal{V}$, and the set of edges by $\mathcal{E}$. The data clustering problem can be expressed as a problem of partitioning a weighted undirected graph $G = (\mathcal{V}, \mathcal{E})$, in which the degree of each vertex i is defined as $d_i = \sum_{j=1}^{N} w_{ij}$.

The goal is to divide the N data points into $K > 1$ clusters. This is equivalent to partitioning $\mathcal{V}$ into K disjoint subsets $\{\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_K\}$, where $\bigcup_{k=1}^{K} \mathcal{V}_k = \mathcal{V}$ and $\mathcal{V}_k \cap \mathcal{V}_{k'} = \emptyset$ for any $k \neq k'$. The volume of each cluster k is defined as $\mathrm{vol}(\mathcal{V}_k) = \sum_{i\in\mathcal{V}_k} d_i$. Intuitively, data clustering aims to group together those data points with high similarities between them. But it is also important to regularize the cluster sizes, as otherwise the algorithm tends to put almost all data points in one giant cluster and leaves other clusters almost empty, leading to the cluster imbalance problem. The above goal can be accomplished by minimizing

$$\mathrm{ncut}(\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_K) = \sum_{k=1}^{K} \frac{\sum_{i\in\mathcal{V}_k} \sum_{j\notin\mathcal{V}_k} w_{ij}}{\mathrm{vol}(\mathcal{V}_k)}. \tag{41}$$

We can rewrite the problem by introducing indicator variable $x_{ik} \in \{0, 1\}$, which equals to 1 if the data point is assigned to cluster k, and equals 0 otherwise. Let $\boldsymbol{x}_k = [x_{1k}, x_{2k}, \ldots, x_{Nk}]^\top$ and $\boldsymbol{W} = [w_{ij}] \in \mathbb{R}^{N\times N}$. We remark that $\boldsymbol{W}$ is typically positive definite (e.g., when the similarities are generated by a Gaussian kernel). It can be shown that the normalized-cut minimization problem boils down to

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{maximize}} \quad & \sum_{k=1}^{K} \frac{\boldsymbol{x}_k^\top \boldsymbol{W} \boldsymbol{x}_k}{\boldsymbol{d}^\top \boldsymbol{x}_k} \\ \text{subject to} \quad & \boldsymbol{x}_k \in \{0, 1\}^N, \quad k = 1, \ldots, K \\ & \sum_{k=1}^{K} x_{ik} = 1, \quad i = 1, \ldots, N, \end{aligned} \tag{42}$$

where $\boldsymbol{d} = [d_1, d_2, \ldots, d_N]^\top$. The constraint $\sum_{k=1}^{K} x_{ik} = 1$

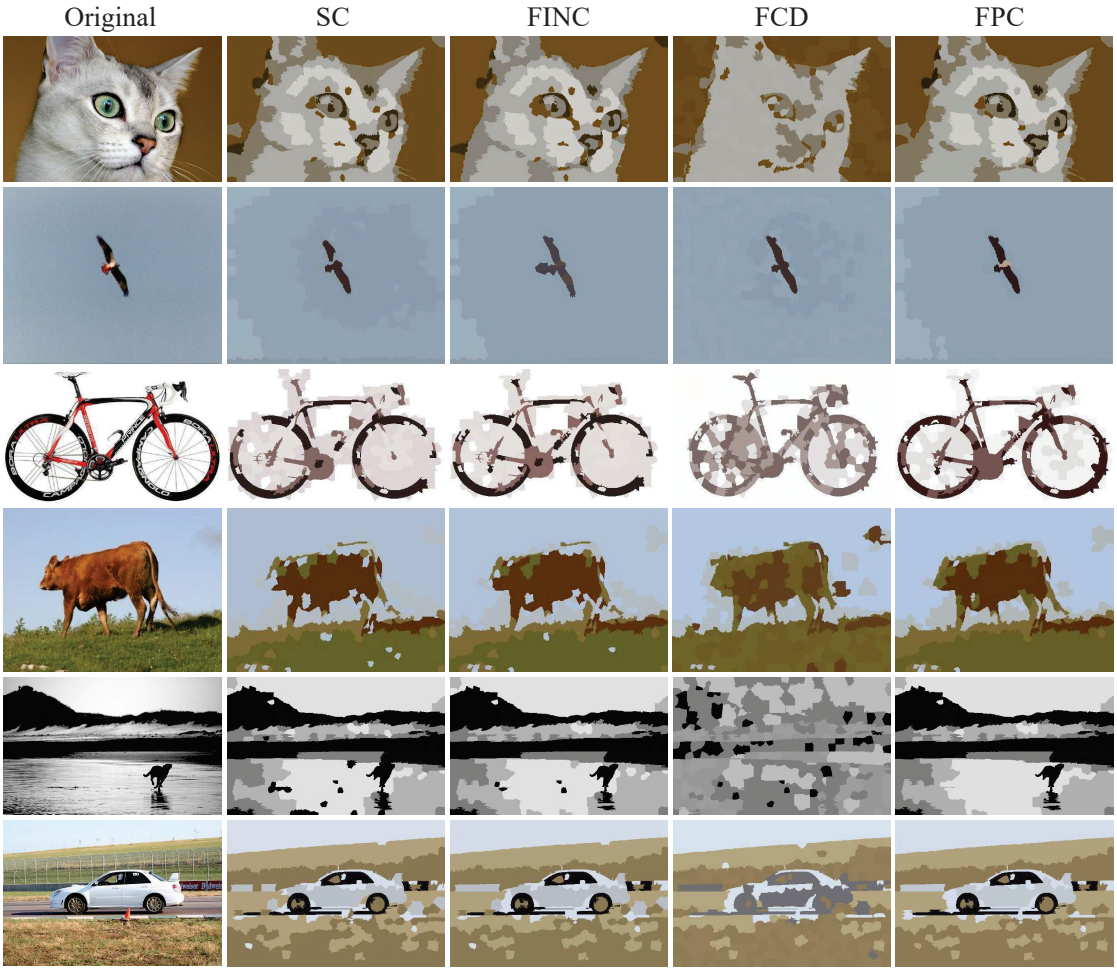


Fig. 7. Image segmentation by the different algorithms based on the normalized-cut problem solving. After decoupling the ratio in the normalized-cut problem by the quadratic transform, we convert the nonlinear discrete optimization to a weighted bipartite matching problem that can be readily solved.

states that each data point i can be assigned to only one cluster. The two constraints ensure that each data point must be assigned to a unique cluster. Because the problem is fractionally structured, it is natural to adopt an FP approach.

The authors of [33] suggest using the scalar quadratic transform to decouple the ratios in the above problem as

$$\begin{aligned} \underset{\boldsymbol{x},\, y}{\text{maximize}} \quad & \sum_{k=1}^{K} \left(2y_k \sqrt{\boldsymbol{x}_k^\top \boldsymbol{W} \boldsymbol{x}_k} - y_k^2 \boldsymbol{d}^\top \boldsymbol{x}_k \right) \\ \text{subject to} \quad & \boldsymbol{x}_k \in \{0,1\}^N, \quad k = 1, \ldots, K \\ & \sum_{k=1}^{K} x_{ik} = 1, \quad i = 1, \ldots, N \\ & y_k \in \mathbb{R}, \quad k = 1, \ldots, K. \end{aligned} \tag{43}$$

Further, [33] proposes to convexify the new optimization problem in $\boldsymbol{x}$ by the Cauchy-Schwarz inequality. The resulting alternating optimization between $\boldsymbol{x}$ and y guarantees monotonically decreasing convergence of $\operatorname{ncut}(\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_K)$.

A better way to solve this problem is to apply the matrix quadratic transform. First, factorize the numerator as $\boldsymbol{x}_k^\top \boldsymbol{W} \boldsymbol{x}_k = \boldsymbol{z}_k \boldsymbol{z}_k^\top$ where $\boldsymbol{z}_k = \boldsymbol{x}_k^\top \boldsymbol{W}^{\frac{1}{2}} \in \mathbb{R}^{1 \times N}$ and $\boldsymbol{W}^{\frac{1}{2}} \in \mathbb{S}_+^{N \times N}$ is the symmetric square root of $\boldsymbol{W}$. We then rewrite $\frac{\boldsymbol{x}_k^\top \boldsymbol{W} \boldsymbol{x}_k}{\boldsymbol{d}^\top \boldsymbol{x}_k}$ as $\operatorname{Tr}(\boldsymbol{z}_k^\top (\boldsymbol{d}^\top \boldsymbol{x}_k)^{-1} \boldsymbol{z}_k)$. At first glance, it seems peculiar to rewrite the scalar ratio in a much more complicated matrix form, but this will pay off soon. By the matrix quadratic transform (39), we convert the original normalized-cut problem to

$$\begin{aligned} \underset{\boldsymbol{x},\, \boldsymbol{y}}{\text{maximize}} \quad & \sum_{k=1}^{K} \operatorname{Tr}\left(2\boldsymbol{y}_k \boldsymbol{x}_k^\top \boldsymbol{W}^{\frac{1}{2}} - \boldsymbol{y}_k \boldsymbol{d}^\top \boldsymbol{x}_k \boldsymbol{y}_k^\top \right) \\ \text{subject to} \quad & \boldsymbol{x}_k \in \{0,1\}^N, \quad k = 1, \ldots, K \\ & \sum_{k=1}^{K} x_{ik} = 1, \quad i = 1, \ldots, N \\ & \boldsymbol{y}_k \in \mathbb{R}^N, \quad k = 1, \ldots, K. \end{aligned} \tag{44}$$

The key observation from [50] is that solving for $\boldsymbol{x}$ in the above new problem under fixed $\boldsymbol{y}$ is a weighted bipartite matching problem that can be solved in closed form as

$$x_{ik}^\star = \begin{cases} 1 & \text{if } k = \arg\max\limits_{k'} \mu_{ik'} \\ 0 & \text{otherwise} \end{cases}, \tag{45}$$

where μ_{ik} is the ith component of $\boldsymbol{\mu}_k = 2\boldsymbol{W}^{\frac{1}{2}} \boldsymbol{y}_k - \boldsymbol{y}_k^\top \boldsymbol{y}_k \boldsymbol{d}$. This alternating optimization between $\boldsymbol{x}$ and $\boldsymbol{y}$ is referred to as

the fractional programming based clustering (FPC) algorithm.

We illustrate the performance advantage of FPC in an image segmentation task. The benchmarks are the spectral clustering (SC) algorithm [51], the fast iterative normalized cut (FINC) algorithm [33], and the fast coordinate descent (FCD) algorithm [52]. We use the Gaussian kernel to generate the similarity matrix, i.e., $w_{ij} = \exp\left(-\|\boldsymbol{v}_i - \boldsymbol{v}_j\|_2^2\right)$, where $\boldsymbol{v}_i$ and $\boldsymbol{v}_i$ are the feature vectors of data points i and j. As shown in Fig. 7, clustering by FPC gives clearer boundaries of the objects than the other methods.

Example 7 (Pilot Signal Design for Channel Estimation): Consider a wireless cellular network consisting of L cells, with one BS and K user terminals per cell. We use i or j to index each cell and its BS; the kth user in cell i is indexed as (i,k). Assume that every BS has N antennas and every user terminal has a single antenna. Let $\boldsymbol{h}_{ijk} \in \mathbb{C}^N$ be the channel from user (j,k) to BS i, and let $\boldsymbol{H}_{ij} = [\boldsymbol{h}_{ij1}, \boldsymbol{h}_{ij2}, \ldots, \boldsymbol{h}_{ijK}]$, with each $\boldsymbol{h}_{ijk}$ modeled as $\boldsymbol{h}_{ijk} = \boldsymbol{g}_{ijk}\sqrt{\beta_{ijk}}$, where $\boldsymbol{g}_{ijk} \in \mathbb{C}^N$ is the unknown small-scale fading coefficient with i.i.d. entries distributed as $\mathcal{CN}(0,1)$, and $\beta_{ijk} \geq 0$ is the large-scale fading coefficient, assumed to be known. Each BS i estimates its $\boldsymbol{H}_{ii}$ based on the uplink pilot signals from the users in the cell. Let $\boldsymbol{s}_{ik} \in \mathbb{C}^\tau$ be a sequence of pilot symbols transmitted from user (i,k), and let $\boldsymbol{S}_i = [\boldsymbol{s}_{i1}, \boldsymbol{s}_{i2}, \ldots, \boldsymbol{s}_{iK}]$. Due to the limited pilot length, the pilot sequences across all the cells cannot all be orthogonal. This results in pilot contamination. We aim to design $S_1, \cdots, S_L$ based on the large-scale fading coefficients β_{ijk} to minimize pilot contamination across the L cells.

The pilot signal received at BS i is

$$\boldsymbol{V}_i = \boldsymbol{H}_{ii}\boldsymbol{S}_i^\top + \sum\nolimits_{j\neq i} \boldsymbol{H}_{ij}\boldsymbol{S}_j^\top + \boldsymbol{Z}_i, \tag{46}$$

where $\boldsymbol{Z}_i \in \mathbb{C}^{N\times\tau}$ is the additive noise with i.i.d. entries distributed as $\mathcal{CN}(0,\sigma^2)$. Let $\hat{\boldsymbol{h}}_{iik}$ be the minimum mean squared error (MMSE) estimate of $\boldsymbol{h}_{iik}$ based on $\boldsymbol{V}_i$. We aim to minimize the sum of mean squared errors (MSEs):

$$\underset{\boldsymbol{S}}{\text{minimize}} \quad \sum_{i=1}^{L}\sum_{k=1}^{K} \mathbb{E}\Big[\|\widehat{\boldsymbol{h}}_{iik} - \boldsymbol{h}_{iik}\|_2^2 \mid \boldsymbol{V}_i\Big]. \tag{47}$$

After a bit of algebra, the MSE minimization problem can be rewritten as

$$\begin{aligned} \underset{\boldsymbol{S}}{\text{maximize}} \quad & \sum_{i=1}^{L} \mathrm{Tr}\big(\boldsymbol{P}_{ii}\boldsymbol{S}_i^{\mathrm{H}}\boldsymbol{D}_i^{-1}\boldsymbol{S}_i\boldsymbol{P}_{ii}\big) \\ \text{subject to} \quad & \|\boldsymbol{s}_{ik}\|_2^2 \leq \rho, \quad \text{for any pair } (i,k), \end{aligned} \tag{48}$$

where $\boldsymbol{P}_{ij} = \mathrm{diag}[\beta_{ij1}, \beta_{ij2}, \ldots, \beta_{ijK}]$, ρ is the power constraint, and $\boldsymbol{D}_i = \sigma^2\boldsymbol{I}_\tau + \sum_{j=1}^{L}\boldsymbol{S}_j\boldsymbol{P}_{ij}\boldsymbol{S}_j^{\mathrm{H}}$. The above problem can be immediately addressed by the matrix quadratic transform, by treating $\sqrt{\boldsymbol{A}_i} = \boldsymbol{S}_i\boldsymbol{P}_{ii}$ and $\boldsymbol{B}_i = \boldsymbol{D}_i$. The resulting pilot design is referred to as the fractional programming pilot (FPP).

We validate the performance of the FP method in a 7-cell wrapped-around network. Each cell comprises a 16-antenna BS and 9 single-antenna user terminals uniformly distributed. The BS-to-BS distance is 1000 meters. Let $\tau = 10$ and let $\rho = 1$. Assume that the background noise is negligible and that $\beta_{ijk} = \varphi_{ijk}/d_{ijk}^3$ where φ_{ijk} is an i.i.d. log-normal random variable according to $\mathcal{N}(0, 8^2)$ and d_{ijk} is the distance between user (j,k) and BS i. Consider two baseline methods: (i) *Orthogonal Method* that fixes a set of 10 orthogonal pilots and allocates a random subset of 9 pilots to users in each cell; (ii) *Random Method* that generates the pilots randomly and independently according to the Gaussian distribution. The orthogonal method is used to initialize the FPP method. The simulation results are shown in Fig. 8. Observe that FFP achieves much smaller MSE overall than the benchmarks.

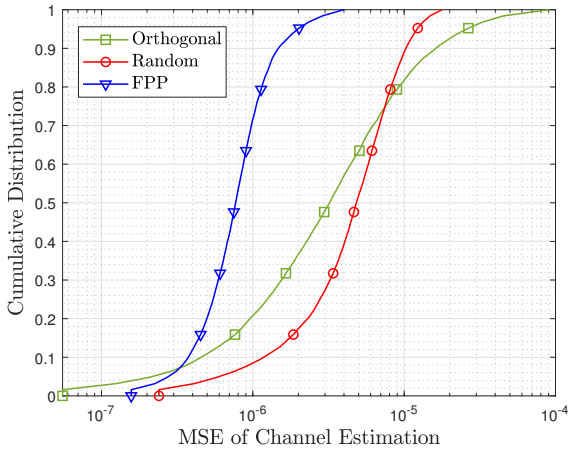


Fig. 8. Cumulative distributions of the MSEs of channel estimation by the different pilot design methods.

VIII. Connections with Various Aspects of Optimization Theory

A. Connection with MM Method

We show that the quadratic transform [5], the Lagrangian dual transform [11], and the AM-GM inequality transform [31] can all be interpreted as an MM method [45], [46]. A brief review of the MM method is as follows. For the primal problem

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad f_o(x), \tag{49}$$

where $f_o(x)$ is typically nonconcave, the MM method constructs a so-called *surrogate function* $g(x|\hat{x})$ conditioned on the parameter $\hat{x}$ that satisfies the following conditions:

$$g(x|\hat{x}) \leq f_o(x) \;\; \forall x \in \mathcal{X} \;\text{ and }\; g(\hat{x}|\hat{x}) = f_o(\hat{x}). \tag{50}$$

Then instead of optimizing x directly in the primal problem, the MM method solves the new problem

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad g(x|\hat{x}), \tag{51}$$

where the condition variable $\hat{x}$ is updated to the previous solution x iteratively. Optimizing x for fixed $\hat{x}$ is referred to as the *maximization* step, while updating $\hat{x}$ for the current x is referred to as the *minorization* step.

It turns out that the entire family of the quadratic transform methods can all be interpreted as MM methods. Let us take

the unified quadratic transform as an example. In this case,

$$f_o(x) = \sum_{i=1}^{n^+} f_i^+ \left(\frac{A_i^+(x)}{B_i^+(x)} \right) + \sum_{i=1}^{n^-} f_i^- \left(\frac{A_i^-(x)}{B_i^-(x)} \right). \quad (52)$$

We now treat the optimal update of y_i in (17) as a function of $\hat{x}$, i.e.,

$$\mathcal{Y}_i(\hat{x}) = \sqrt{A_i^+(\hat{x})}/B_i^+(\hat{x}), \quad (53)$$

and replace each y_i with $\mathcal{Y}_i(\hat{x})$ in the new optimization objective in (25); likewise, we replace $\tilde{y}_i$ by $\widetilde{\mathcal{Y}}_i(\hat{x})$ where

$$\widetilde{\mathcal{Y}}_i(\hat{x}) = \sqrt{B_i^-(\hat{x})}/A_i^-(\hat{x}). \quad (54)$$

This gives a conditional function

$$g(x|\hat{x}) = \sum_{i=1}^{n^+} f_i^+ \Big(2\mathcal{Y}_i(\hat{x}) \sqrt{A_i^+(x)} - \mathcal{Y}_i^2(\hat{x}) B_i^+(x) \Big) + \sum_{i=1}^{n^-} f_i^- \Big(\big(\big[2\widetilde{\mathcal{Y}}_i(\hat{x}) \sqrt{A_i^-(x)} - \widetilde{\mathcal{Y}}_i^2(\hat{x}) B_i^-(x) \big]_+ \big)^{-1} \Big). \quad (55)$$

It can be shown that the above $g(x|\hat{x})$ satisfies the defining properties of MM in (50), so it is a valid surrogate function. Thus, for the alternating optimization between x and y by the unified quadratic transform, we can view the update of x as the maximization step, and the update of y as the minorization step.

Likewise, for the channel capacity maximization problem

$$f_o(x) = \sum_{i=1}^{n} w_i \log \left(1 + \frac{A_i(x)}{B_i(x)} \right), \quad (56)$$

the Lagrangian dual transform in (31) can be interpreted as constructing a surrogate function

$$g(x|\hat{x}) = \sum_{i=1}^{n} w_i \Big(\log \big(1 + \Gamma_i(\hat{x}) \big) + \frac{(1+\Gamma_i(x))A_i(x)}{A_i(x) + B_i(x)} - \Gamma_i(\hat{x}) \Big), \quad (57)$$

where $\Gamma_i(\hat{x}) = \frac{A_i(\hat{x})}{B_i(\hat{x})}$, so it also belongs to the MM family.

Moreover, the AM-GM inequality transform from [31] for the sum-of-ratios min problem (18) can be interpreted as an MM method as well. In this case, we have

$$f_o(x) = -\sum_{i=1}^{n} \frac{A_i(x)}{B_i(x)}. \quad (58)$$

As before, we treat the optimal update of the auxiliary variable y_i in (22) as a function

$$\mathcal{Y}_i(\hat{x}) = \frac{2}{A_i(\hat{x}) B_i(\hat{x})}, \quad (59)$$

and then replace each y_i with $\mathcal{Y}_i(\hat{x})$ in the new objective in (21) to obtain

$$g(x|\hat{x}) = -\sum_{n=1}^{N} w_i \left(\mathcal{Y}_i(\hat{x}) A_i^2(x) + \frac{1}{4\mathcal{Y}_i(\hat{x}) B_i^2(x)} \right). \quad (60)$$

The fact that the above $g(x|\hat{x})$ satisfies the defining properties of MM (50) follows from the AM-GM inequality:

$$\mathcal{Y}_i(\hat{x}) A_i^2(x) + \frac{1}{4\mathcal{Y}_i(\hat{x}) B_i^2(x)} \geq \frac{A_i(x)}{B_i(x)}, \quad (61)$$

where the equality holds if and only if $\mathcal{Y}_i(\hat{x}) A_i^2(x) = 1/(4\mathcal{Y}_i(\hat{x}) B_i^2(x))$, i.e., when $\mathcal{Y}_i(\hat{x})$ equals y_i in (22).

The MM interpretation of these FP methods brings two benefits. First, it justifies the composition of the quadratic transform and other methods, e.g., the use of the Lagrangian dual transform in conjunction with the quadratic transform for power control as illustrated in one preceding example. This is because the composition is still an MM method. (In contrast, because Dinkelbach's method does not belong to the MM family, its combination with the other FP methods cannot be as easily justified.)

Second, with the MM interpretation at hand, we can readily derive convergence conditions for the FP methods. The MM interpretation can further enable the convergence rate analysis as discussed later; intuitively, the convergence rate depends on how tight $g(x|\hat{x})$ approximates $f_o(x)$. We defer the detailed discussion to the later part of this article.

Since the quadratic transform can be interpreted as an MM method, it would be tempting to consider alternative MM methods by choosing other surrogate functions. But it is not always easy to determine the best surrogate function in practice. For instance, one may suggest using the first-order Taylor expansion, but the resulting linear surrogate function is less tight than the quadratic surrogate function in the quadratic transform; the tightness of various surrogate functions for MM is shown in Fig. 12 later in the article. Moreover, one may suggest using the second-order Taylor expansion to construct an alternative quadratic surrogate function, but its application is more limited than the quadratic transform, because it requires the ratio function to be continuously differentiable with a fixed Lipschitz constant; in contrast, the quadratic transform only requires the basic assumptions for FP, i.e., each numerator is nonnegative while each denominator is strictly positive. Further, it is possible to construct an even tighter surrogate function than the quadratic form for some specific cases, but this surrogate function may not be easy to optimize, not to mention potential generalizability issues.

B. Connection with Gradient Projection: Accelerating the Quadratic Transform

We now focus attention on the following type of matrix ratio:

$$M_i = \big(\boldsymbol{A}_i \boldsymbol{x}_i \big)^{\mathrm{H}} \left(\sum_{j=1}^{n} \boldsymbol{B}_{ij} \boldsymbol{x}_j \boldsymbol{x}_j^{\mathrm{H}} \boldsymbol{B}_{ij}^{\mathrm{H}} \right)^{-1} \big(\boldsymbol{A}_i \boldsymbol{x}_i \big), \quad (62)$$

where $\boldsymbol{x}_j \in \mathbb{C}^m$, $\boldsymbol{A}_i \in \mathbb{C}^{\ell \times m}$, and $\boldsymbol{B}_{ij} \in \mathbb{C}^{\ell \times m}$, for $i, j = 1, \ldots, n$. Typically, $\ell \ll m$. The above ratio term is of particular interest because many important metrics in practice can be written in this form, e.g., the CRB, the Fisher information, the SINR, and the normalized cut. Consider the

Algorithm 1: Different approaches to solving problem (63).

initialize $\boldsymbol{x}$ to a feasible value;
repeat
  update each $\boldsymbol{y}_i = \big(\sum_{j=1}^n \boldsymbol{B}_{ij}\boldsymbol{x}_j\boldsymbol{x}_j^{\mathrm{H}}\boldsymbol{B}_{ij}^{\mathrm{H}}\big)^{-1}\big(\boldsymbol{A}_i\boldsymbol{x}_i\big)$;
  switch *the choice of transform* **do**
    case *basic quadratic transform* **do**
      update each $\boldsymbol{x}_i$ according to (65);
    end
    case *nonhomogeneous quadratic transform* **do**
      update each $\boldsymbol{x}_i$ according to (68);
    end
    case *extrapolated quadratic transform* **do**
      update each $\boldsymbol{\nu}_i$ according to (71);
      update each $\boldsymbol{x}_i$ according to (72);
    end
  end
until *the value of* $f_o(\boldsymbol{x})$ *converges*;

sum-of-weighted-matrix-ratios problem involving a total of n such matrix ratios:

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{maximize}} \quad & f_o(\boldsymbol{x}) := \sum_{i=1}^n w_i M_i \\ \text{subject to} \quad & \boldsymbol{x}_i \in \mathcal{X}_i, \quad i = 1, \ldots, n, \end{aligned} \tag{63}$$

where each $w_i > 0$ is a strictly positive weight and each $\mathcal{X}_i$ is a nonempty convex constraint set on $\boldsymbol{x}_i$. By the quadratic transform [5], the original optimization objective $f_o(\boldsymbol{x})$ can be recast to

$$f_q(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^n w_i \left(2\Re\{\boldsymbol{x}_i^{\mathrm{H}}\boldsymbol{A}_i^{\mathrm{H}}\boldsymbol{y}_i\} - \sum_{j=1}^n \boldsymbol{y}_i^{\mathrm{H}}\boldsymbol{B}_{ij}\boldsymbol{x}_j\boldsymbol{x}_j^{\mathrm{H}}\boldsymbol{B}_{ij}^{\mathrm{H}}\boldsymbol{y}_i \right) \tag{64}$$

with the auxiliary variable $\boldsymbol{y}_i \in \mathbb{C}^\ell$ introduced for each matrix ratio. When $\boldsymbol{x}$ is fixed, each $\boldsymbol{y}_i$ is optimally determined in closed form. When $\boldsymbol{y}$ is fixed, each $\boldsymbol{x}_i$ is optimally determined as

$$\boldsymbol{x}_i^\star = \arg\min_{\boldsymbol{x}_i \in \mathcal{X}_i} \big\| \boldsymbol{D}_i^{\frac{1}{2}} \big(\boldsymbol{x}_i - w_i \boldsymbol{D}_i^{-1}\boldsymbol{A}_i^{\mathrm{H}}\boldsymbol{y}_i\big) \big\|_2, \tag{65}$$

where $\boldsymbol{D}_i = \sum_{j=1}^n w_j \boldsymbol{B}_{ji}^{\mathrm{H}}\boldsymbol{y}_j\boldsymbol{y}_j^{\mathrm{H}}\boldsymbol{B}_{ji}$. A practical issue here is that the matrix inverse in (65) can be quite costly when $\boldsymbol{D}_i$ is a large matrix.

The so-called *nonhomogeneous quadratic transform* [13] aims to eliminate the matrix inverse operation. The main idea is to construct a lower bound on $f_q(\boldsymbol{x}, \boldsymbol{y})$ as

$$\begin{aligned} f_t(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \sum_{i=1}^n \Big(& 2\Re\big\{w_i \boldsymbol{x}_i^{\mathrm{H}}\boldsymbol{A}_i^{\mathrm{H}}\boldsymbol{y}_i + \boldsymbol{x}_i^{\mathrm{H}}(\lambda_i \boldsymbol{I} - \boldsymbol{D}_i)\boldsymbol{z}_i\big\} \\ & + \boldsymbol{z}_i^{\mathrm{H}}(\boldsymbol{D}_i - \lambda_i \boldsymbol{I})\boldsymbol{z}_i - \lambda_i \boldsymbol{x}_i^{\mathrm{H}}\boldsymbol{x}_i \Big), \end{aligned} \tag{66}$$

where $\lambda_i \geq \lambda_{\max}(\boldsymbol{D}_i)$, e.g., $\lambda_i = \|\boldsymbol{D}_i\|_{\mathrm{F}}$. We have $f_q(\boldsymbol{x}, \boldsymbol{y}) \geq f_t(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ in general, where the equality holds if $\boldsymbol{z} = \boldsymbol{x}$. As a consequence, the problem of maximizing $f_q(\boldsymbol{x}, \boldsymbol{y})$ over $\boldsymbol{x}$ and $\boldsymbol{y}$ is equivalent to the problem of maximizing $f_t(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ over $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$. We then consider optimizing $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ iteratively in $f_t(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$. When $\boldsymbol{y}$ and $\boldsymbol{x}$ are both held fixed, the optimal update of $\boldsymbol{z}$ is $\boldsymbol{z}_i^\star = \boldsymbol{x}_i$. When $\boldsymbol{z}$ and $\boldsymbol{x}$ are both fixed, the optimal update of each $\boldsymbol{y}_i$ is

$$\boldsymbol{y}_i = \left(\sum_{j=1}^n \boldsymbol{B}_{ij}\boldsymbol{x}_j\boldsymbol{x}_j^{\mathrm{H}}\boldsymbol{B}_{ij}^{\mathrm{H}} \right)^{-1} \big(\boldsymbol{A}_i\boldsymbol{x}_i\big). \tag{67}$$

Next, when $\boldsymbol{y}$ and $\boldsymbol{z}$ are both held fixed, the optimal $\boldsymbol{x}_i$ is given by

$$\boldsymbol{x}_i^\star = \arg\min_{\boldsymbol{x}_i \in \mathcal{X}_i} \big\| \lambda_i \boldsymbol{x}_i - \omega_i \boldsymbol{A}_i^{\mathrm{H}}\boldsymbol{y}_i - \big(\lambda_i \boldsymbol{I} - \boldsymbol{D}_i\big)\boldsymbol{z}_i \big\|_2. \tag{68}$$

Observe that the computation of the matrix inverse of the $m \times m$ matrix $\boldsymbol{D}_i$ is no longer required. Instead, the only matrix inverse needed is in the update of $\boldsymbol{y}$ in (67). Note that the matrix inverse in (67) is typically of a much smaller dimension, i.e., it is $\ell \times \ell$ instead of $m \times m$ as in (65).

Interestingly, the nonhomogeneous quadratic transform has an intimate connection with gradient projection. We use the superscript $k \in \{1, 2, \ldots\}$ to index the iteration, and let $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ be cyclically updated as $\boldsymbol{x}^{(0)} \to \cdots \to \boldsymbol{z}^{(k)} \to \boldsymbol{y}^{(k)} \to \boldsymbol{x}^{(k)} \to \boldsymbol{z}^{(k+1)} \to \cdots$. We can then rewrite the optimal update of $\boldsymbol{x}_i$ in (68) as the Euclidean projection onto the surface of a sphere:

$$\boldsymbol{x}_i^k = \mathcal{P}_{\mathcal{X}_i}\left(\boldsymbol{z}_i^{(k)} + \frac{1}{\lambda_i}\Big(w_i \boldsymbol{A}_i^{\mathrm{H}}\boldsymbol{y}_i^{(k)} - \boldsymbol{D}_i \boldsymbol{z}_i^{(k)} \Big) \right), \tag{69}$$

which can be further recognized as gradient projection [14]:

$$\boldsymbol{x}_i^{(k)} = \mathcal{P}_{\mathcal{X}_i}\left(\boldsymbol{x}_i^{(k-1)} + \frac{1}{\lambda_i} \cdot \frac{\partial f_o(\boldsymbol{x}^{(k-1)})}{\partial \boldsymbol{x}_i} \right). \tag{70}$$

We now compare the basic quadratic transform and the nonhomogeneous quadratic transform from a geometric perspective. As shown in Fig. 9, the update of $\boldsymbol{x}_i$ in (65) by the basic quadratic transform can be interpreted as the projection onto the surface of an ellipsoid, while the update of $\boldsymbol{x}_i$ in (68) by the nonhomogeneous quadratic transform can be interpreted as the projection onto the surface of a sphere. Projection onto an ellipsoid is in general much more costly to do than projection onto a sphere in a high-dimensional space.

The fact that the nonhomogeneous quadratic transform amounts to a gradient projection method motivates us to investigate the possibility of incorporating Nesterov's extrapolation scheme [53] into FP in order to accelerate its convergence.

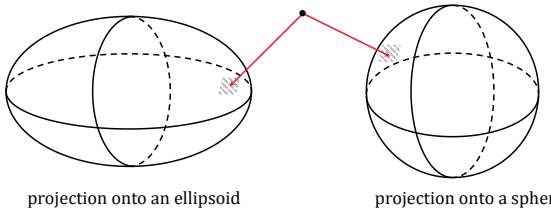


Fig. 9. The basic quadratic transform can be interpreted as the projection onto an ellipsoid in (65), while the nonhomogeneous quadratic transform can be interpreted as the projection onto a sphere in (68). The nonhomogeneous quadratic transform eliminates the matrix inversion when solving the matrix-ratio problem (63).

Specifically, following the *heavy-ball* intuition, we extrapolate each $\boldsymbol{x}_i$ along the direction of the difference between the preceding two iterates before the gradient projection, i.e.,

$$\boldsymbol{\nu}_i^{(k-1)} = \boldsymbol{x}_i^{(k-1)} + \eta_{k-1}(\boldsymbol{x}_i^{(k-1)} - \boldsymbol{x}_i^{(k-2)}), \tag{71}$$

$$\boldsymbol{x}_i^{(k)} = \mathcal{P}_{\mathcal{X}}\left(\boldsymbol{\nu}_i^{(k-1)} + \frac{1}{\lambda_i} \cdot \frac{\partial f_o(\boldsymbol{\nu}^{(k-1)})}{\partial \boldsymbol{x}_i}\right), \tag{72}$$

where the extrapolation stepsize $\eta_k = \max\left\{\frac{k-2}{k+1}, 0\right\}$ and the starting point $\boldsymbol{x}^{(-1)} = \boldsymbol{x}^{(0)}$, as in [53]. We refer to this algorithm as the *extrapolated quadratic transform*. Moreover, a recent work [54] suggests that the extrapolation scheme can be incorporated into the MM method, so it is possible to accelerate the quadratic transform directly without using the nonhomogeneous quadratic transform. Algorithm 1 summarizes the basic, the nonhomogeneous, and the extrapolated quadratic transforms.

C. Connection with WMMSE Algorithm

We now discuss the connections of FP to specific optimization techniques in wireless communications. The WMMSE algorithm [55], [56] is a well-known algorithm for computing the optimal beamformers in a multi-cell MIMO transmission scenario. This section aims to show that the WMMSE algorithm is a special case of the FP method, and moreover such connection enables further improvement of the WMMSE algorithm.

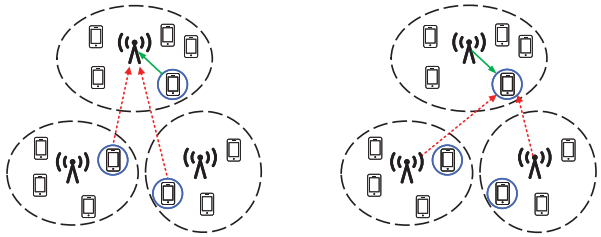
Consider a wireless cellular network with L cells and assume that there are K downlink users in each cell. Assume also that each BS has M transmit antennas and each user has N receive antennas. The kth user in the ith cell is indexed as (i,k). Denote by $\boldsymbol{H}_{ik,j} \in \mathbb{C}^{N\times M}$ the channel from user i to the BS, σ^2 the background noise power, and $\boldsymbol{V}_{ik} \in \mathbb{C}^{M\times d}$ the transmit beamformer of BS i intended for user (i,k), where d is the number of data streams. Let $\boldsymbol{B}_{ik} = \sigma^2\boldsymbol{I} + \sum_{(j,q)\neq(i,k)} \boldsymbol{H}_{ik,j}\boldsymbol{V}_{jq}\boldsymbol{V}_{jq}^{\mathrm{H}}\boldsymbol{H}_{ik,j}^{\mathrm{H}}$ and let $\sqrt{\boldsymbol{A}_{ik}} = \boldsymbol{H}_{ik,i}\boldsymbol{V}_{ik}$. The sum-of-weighted-rates maximization problem aims to maximize the objective of

$$f_o(\boldsymbol{V}) = \sum_{(i,k)} w_{ik} \log\det\left(\boldsymbol{I} + \sqrt{\boldsymbol{A}_{ik}}^{\mathrm{H}}\boldsymbol{B}_{ik}^{-1}\sqrt{\boldsymbol{A}_{ik}}\right), \tag{73}$$

where the weight $w_{ik} > 0$ reflects the priority of user (i,k). While the previous works [55], [56] obtain the WMMSE algorithm from the relationship between SINR and MMSE, we rederive WMMSE by purely using FP. First, by the matrix extension of the Lagrangian dual transform [12], the original objective function can be converted to

$$f_r(\boldsymbol{V},\boldsymbol{\Gamma}) = \sum_{(i,k)} \Big(w_{ik}\log\det(\boldsymbol{I}+\boldsymbol{\Gamma}_{ik}) - w_{ik}\mathrm{Tr}\big(\boldsymbol{\Gamma}_{ik}\big) + w_{ik}\mathrm{Tr}\big((\boldsymbol{I}+\boldsymbol{\Gamma}_{ik})\sqrt{\boldsymbol{A}_{ik}}^{\mathrm{H}}(\boldsymbol{A}_{ik}+\boldsymbol{B}_{ik})^{-1}\sqrt{\boldsymbol{A}_{ik}}\big)\Big). \tag{74}$$

When $\boldsymbol{V}$ is fixed, each $\boldsymbol{\Gamma}_{ik}$ is optimally updated as $\boldsymbol{\Gamma}_{ik}^{\star} = \sqrt{\boldsymbol{A}_{ik}}^{\mathrm{H}}\boldsymbol{B}_{ik}^{-1}\sqrt{\boldsymbol{A}_{ik}}$. It remains to optimize $\boldsymbol{V}$ for fixed $\boldsymbol{\Gamma}$. Maximizing $f_r(\boldsymbol{V},\boldsymbol{\Gamma})$ over $\boldsymbol{V}$ for fixed $\boldsymbol{\Gamma}$ can be recognized as a sum-of-matrix-ratios problem. The quadratic transform can be performed in different ways: one particular way leads to WMMSE [55], [56], while another leads to the FPLinQ algorithm in [11], [12]. Thus, WMMSE can be interpreted as a particular case of the quadratic transform method.



(a) Uplink multicell network (b) Downlink multicell network

Fig. 10. Interference pattern depends on the user scheduling in the neighboring cells in the uplink, but not so in the downlink. Here, the solid lines represent the desired signal; the dashed lines represent the interfering signal; the scheduled user terminal in each cell is circled. The quadratic transform [11] is more powerful than WMMSE [55], [56], when the problem includes the discrete optimization of uplink user scheduling.

We now show how to obtain the WMMSE algorithm from the quadratic transform. If we treat $f_r(\boldsymbol{V},\boldsymbol{\Gamma})$ as a weighted sum of the following matrix ratios

$$\underbrace{(\boldsymbol{A}_{ik})}_{\text{numerator}}\ \underbrace{(\boldsymbol{A}_{ik}+\boldsymbol{B}_{ik})^{-1}}_{\text{denominator}}$$

and use the quadratic transform to decouple each matrix ratio, then $f_r(\boldsymbol{V},\boldsymbol{\Gamma})$ can be further recast as

$$f_q^{(\mathrm{I})}(\boldsymbol{V},\boldsymbol{\Gamma},\boldsymbol{Y}) = \sum_{(i,k)} \Big(w_{ik}\log\det(\boldsymbol{I}+\boldsymbol{\Gamma}_{ik}) - w_{ik}\mathrm{Tr}\big(\boldsymbol{\Gamma}_{ik}\big) + w_{ik}\mathrm{Tr}\big((\boldsymbol{I}+\boldsymbol{\Gamma}_{ik})(\sqrt{\boldsymbol{A}_{ik}}^{\mathrm{H}}\boldsymbol{Y}_{ik} + \boldsymbol{Y}_{ik}^{\mathrm{H}}\sqrt{\boldsymbol{A}_{ik}} - \boldsymbol{Y}_{ik}^{\mathrm{H}}\boldsymbol{B}'_{ik}\boldsymbol{Y}_{ik})\big)\Big) \tag{75}$$

with $\boldsymbol{B}'_{ik} = \boldsymbol{A}_{ik} + \boldsymbol{B}_{ik}$. Optimizing $(\boldsymbol{V},\boldsymbol{\Gamma},\boldsymbol{Y})$ iteratively in the above new objective function gives rise exactly to the WMMSE algorithm [55], [56].

There are also other ways of using the quadratic transform to decouple the matrix ratios. For instance, we could have alternatively regarded $f_r(\boldsymbol{V},\boldsymbol{\Gamma})$ as a sum of following matrix ratios:

$$\underbrace{(w_{ik}(\boldsymbol{I}+\boldsymbol{\Gamma}_{ik})\boldsymbol{A}_{ik})}_{\text{numerator}}\ \underbrace{(\boldsymbol{A}_{ik}+\boldsymbol{B}_{ik})^{-1}}_{\text{denominator}}.$$

Differing from the preceding matrix ratio of the WMMSE case, the above matrix includes the weight w_{ik} and the auxiliary variable $\boldsymbol{I}+\Gamma_{ik}$ in the numerator. After decoupling the above matrix ratio by the quadratic transform, we recast $f_r(\boldsymbol{V},\boldsymbol{\Gamma})$ into a different objective function:

$$f_q^{(\mathrm{II})}(\boldsymbol{V},\boldsymbol{\Gamma},\boldsymbol{Y}) = \sum_{(i,k)} \Big(w_{ik}\log\det(\boldsymbol{I}+\boldsymbol{\Gamma}_{ik}) - w_{ik}\mathrm{Tr}\big(\boldsymbol{\Gamma}_{ik}\big) + \sqrt{\boldsymbol{A}'_{ik}}^{\mathrm{H}}\boldsymbol{Y}_{ik} + \boldsymbol{Y}_{ik}^{\mathrm{H}}\sqrt{\boldsymbol{A}'_{ik}} - \boldsymbol{Y}_{ik}^{\mathrm{H}}(\boldsymbol{A}_{ik}+\boldsymbol{B}_{ik})\boldsymbol{Y}_{ik}\Big) \tag{76}$$

with $\sqrt{\boldsymbol{A}'_{ik}} = \sqrt{w_i}(\boldsymbol{I}+\boldsymbol{\Gamma}_i)^{\frac{1}{2}}\sqrt{\boldsymbol{A}_{ik}}$. Again, we consider optimizing $(\boldsymbol{V},\boldsymbol{\Gamma},\boldsymbol{Y})$ in an iterative fashion. This gives rise to a beamforming algorithm called FPLinQ in [11], [12]. We

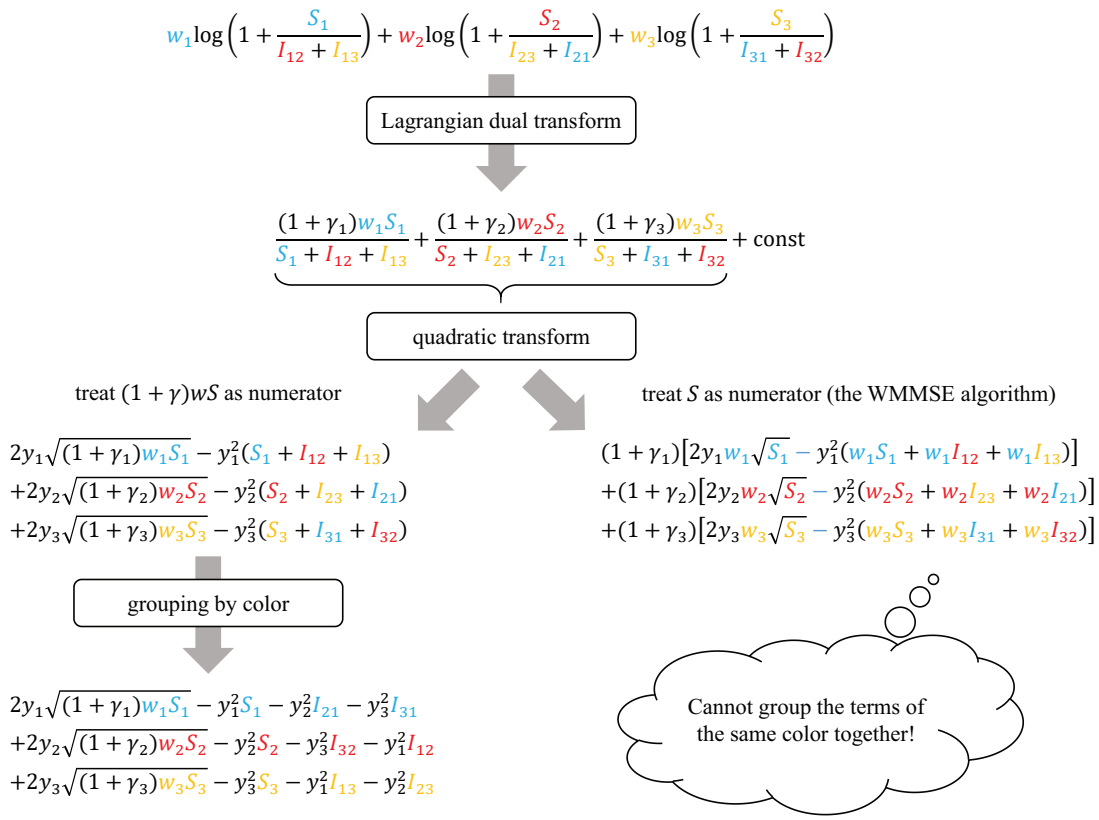


Fig. 11. Consider uplink scheduling for three SISO cells. Let w_i (the rate weight of the scheduled user), S_i (the received signal power), and I_{ij} (the interference from cell j to cell i) be colored by their respective cell index i. After the Lagrangian dual transform, the optimization objective is transformed into a sum-of-weighted-ratios objective plus a constant. In the next step, if we apply the quadratic transform as in FPLinQ [11], then the terms in the new problem can be grouped according to their colors, so that the scheduling problem becomes a bipartite matching problem. In contrast, if we follow the WMMSE algorithm [55], [56], then the variables with distinct colors are coupled, making the discrete optimization problem difficult to tackle.

remark that the above results can be easily extended to more complicated networks. For instance, for the cell-free network [19] wherein each downlink user is served by multiple BSs, the matrix ratio becomes $\big(\sum_{i\in\mathcal{N}_k}\sqrt{\boldsymbol{A}_{ik}}\big)^{\mathrm{H}}\boldsymbol{B}_k^{-1}\big(\sum_{i\in\mathcal{N}_k}\sqrt{\boldsymbol{A}_{ik}}\big)$, where $\mathcal{N}_k$ is the set of BSs assigned to user k, $\sqrt{\boldsymbol{A}_{ik}}$ is the precoded signal sent from BS i intended for user k, and $\boldsymbol{B}_k$ is the interference-plus-noise covariance matrix at user k. The resulting FP problem is mathematically the same as the cellular network discussed above.

Regardless of how the matrix ratio is decoupled by the quadratic transform, the FP method always introduces some auxiliary variables and then optimizes them along with the beamforming variable iteratively. This approach enjoys two advantages: (i) each iterative update can be done in closed form; (ii) it guarantees convergence to a stationary point of the beamforming problem (as shown earlier due to the connection to the MM theory). But are these distinct FP methods (e.g., WMMSE and FPLinQ) equally good? Does it matter in algorithm design which way to decouple the ratio?

The following example sheds light on the above questions. Let us shift attention to the uplink network. To ease notation, consider the single-input single-output (SISO) case, but the following discussion can be generalized to the MIMO case as in [11], [12]. We incorporate uplink scheduling into the problem formulation, i.e., in each time-slot, we need to select one user in each cell for uplink transmission with an aim of maximizing a utility function of long-term average rates across all the users. It is worth noting that the uplink scheduling problem is more challenging than the downlink, as illustrated in Fig. 10. We can carry out the quadratic transform in different ways. If we decouple the ratios as in $f_q^{(\mathrm{I})}$, then the ratio decoupling does not make the scheduling problem much easier since the discrete scheduling variables are still coupled together in the new problem. But if we instead decouple the ratios as in $f_q^{(\mathrm{II})}$, then the discrete variables can be decoupled after the ratio decoupling, so the scheduling problem can be efficiently solved by the standard bipartite matching method. Fig. 11 gives a detailed exposition of the advantage of FPLinQ over WMMSE for the uplink scheduling problem.

D. Connection with Fixed-Point Iteration

We now revisit the earlier example of power control for interfering links, with an aim to connect the FP method with the fixed-point iteration based power control method [57]–[59]. We use $f_o(p)$ to denote the sum-of-weighted-rates objective of the power control problem, i.e.,

$$f_o(p)=\sum_{i=1}^{K}w_i\log\left(1+\frac{|h_{ii}|^2p_i}{\sum_{j\neq i}|h_{ij}|^2p_j+\sigma^2}\right). \tag{77}$$

For ease of discussion, we ignore the total power constraint. Fixed-point iteration is an approach to the power control problem in the literature [57]–[59] based on rewriting the first-order equation $\partial f_o(p)/\partial p_i = 0$ as $p_i = G_i(p)$ whereby p_i is isolated on the left-hand side and is iteratively updated by the previous p_i on the other side, i.e.,

$$p_i^{(t+1)} = G_i(p^{(t)}), \tag{78}$$

where the superscript t or $t+1$ is the iteration index. Note that $\partial f_o(p)/\partial p_i = 0$ can be rewritten as $p_i = G_i(p)$ in different ways, each leading to a different fixed-point iteration. If the iterative update of p_i converges for every i, then the first-order condition must hold after convergence and thus we end up with a stationary-point solution.

However, it is not an easy task to prove the convergence of a fixed-point iteration. A variety of fixed-point iteration methods have been proposed in [57]–[59] in pursuit of convergence. For instance, the authors of [58] suggest rewriting $\partial f_o(p)/\partial p_i = 0$ as

$$p_i = \left(\frac{w_i\Gamma_i}{1+\Gamma_i}\right)\left(\sum_{j=1,j\neq i}^{K}\frac{w_j\Gamma_i^2|h_{ji}|^2}{(1+\Gamma_i)|h_{jj}|^2p_j}\right)^{-1}, \tag{79}$$

with

$$\Gamma_i = \frac{|h_{ii}|^2p_i}{\sum_{j\neq i}|h_{ij}|^2p_j+\sigma^2}. \tag{80}$$

It is shown in [58] that the above fixed-point iteration can guarantee convergence provided that the initial value of every $\Gamma_i(p^{(0)})$ is sufficiently high, by exploiting the *standard function properties* [59]. However, finding a fixed-point iteration with provable convergence in general remains difficult.

Through reverse engineering as shown in [5], the (closed-form) FP method for the power control problem as discussed in the earlier example can be interpreted as a fixed-point iteration with $\partial f_o(p)/\partial p_i = 0$ rewritten as

$$p_i = \left(\frac{w_i^2\Gamma_i^2}{p_i}\right)\left(\sum_{j=1}^{K}\frac{w_j\Gamma_i^2|h_{ji}|^2}{(1+\Gamma_i)|h_{jj}|^2p_j}\right)^{-2}. \tag{81}$$

The convergence of this fixed-point iteration now follows directly from the convergence of the FP method—which is verified by the MM theory as discussed earlier in the article.

IX. Convergence Analysis

This section presents two main results on the convergence of the FP algorithms. First, we show that the iterative optimization by the quadratic transform has provable convergence to a stationary-point solution of the original problem, under a weaker condition than that for the block coordinate descent (BCD) method. Second, focusing on the sum-of-weighted-matrix-ratios problem (63), we examine the rate of convergence for the different quadratic transform methods as summarized in Algorithm 1.

The analysis relies on the MM theory [45], [46]. First of all, it is easy to verify a composition property for the surrogate function in (50): if $g(x|\hat{x})$ is a surrogate function of $f_o(x)$ while $h(x|\hat{x})$ is a surrogate function of $g(x|x)$, then $h(x|\hat{x})$ must be a surrogate function of $f_o(x)$. We have already shown that the quadratic transform (including all its extensions) can be interpreted as MM. The Lagrangian dual transform in (31) can also be interpreted as MM. Thus, their composition, as considered in the example of power control for interfering links, also amounts to an MM method. Likewise, because the nonhomogenous bound in [46] in essence is about constructing a surrogate function, its use in conjunction with the quadratic transform, namely the nonhomogeneous quadratic transform, boils down to an MM method too. Thus, by the MM theory, all these FP methods guarantee that the value of the maximization (resp. minimization) objective function of the original FP problem is nondecreasing (resp. nonincreasing) after each iteration. Further, they converge to a stationary point of the original FP problem under certain mild conditions as stated in [45], [46].

For instance, the alternating optimization between x and y in (16) guarantees convergence to a stationary point of the original problem (14) provided that each $A_i(x)$ is differentiable and concave, each $B_i(x)$ is differentiable and convex, and $\mathcal{X}$ is a convex set. Without using the MM interpretation, the alternating optimization can only be viewed as the BCD method, but the convergence condition for BCD is stronger, requiring each iterate to be uniquely solvable [60]. As such, $A_i(x)$ being concave and $B_i(x)$ being convex in (14) are no longer enough for convergence; it now requires strict concavity and strict convexity.

We further analyze how fast the different quadratic transform methods converge for the problem instance (63) by considering the convergence of the function value versus the number of iterates. Consider the basic quadratic transform, the nonhomogeneous quadratic transform, and the extrapolated quadratic transform. Assume that the Hessian of the original objective function $f_o(\boldsymbol{x})$ is L-Lipschitz continuous. To make the analysis tractable, we further require the constraint set to be a small neighborhood of a local maximum $\boldsymbol{x}^*$, so that the distance between $\boldsymbol{x}^*$ and the kth-iteration solution $\boldsymbol{x}^{(k)}$ is bounded, i.e., $\|\boldsymbol{x}-\boldsymbol{x}^*\|_2 \leq R$ for all k. We then analyze the local convergence behavior assuming that the radius R is sufficiently small. Recall that both the basic quadratic transform and nonhomogeneous quadratic transform are equivalent to constructing a surrogate function for $f_o(\boldsymbol{x})$. Let $g(\boldsymbol{x}|\hat{\boldsymbol{x}})$ be the surrogate function due to either the basic quadratic transform or nonhomogeneous quadratic transform; define the gap between the original optimization objective $f_o(\boldsymbol{x})$ and the surrogate function $g(\boldsymbol{x}|\hat{\boldsymbol{x}})$ to be $\delta(\boldsymbol{x}|\hat{\boldsymbol{x}}) = f_o(\boldsymbol{x}) - g(\boldsymbol{x}|\hat{\boldsymbol{x}})$. Moreover, let the maximum eigenvalue of the Hessian of $\delta(\boldsymbol{x}|\hat{\boldsymbol{x}})$ be Λ. According to [14], the basic quadratic transform and nonhomogeneous quadratic transform have the following convergence rate:

$$f_o(\boldsymbol{x}^*) - f_o(\boldsymbol{x}^{(k)}) \leq \frac{2\Lambda R^2 + 2LR^3/3}{k+3}, \quad \text{for } k \geq 2, \tag{82}$$

where the parameter $\Lambda \geq 0$ differs for the two transforms. It can be shown that Λ of the basic quadratic transform is smaller than that of the nonhomogeneous quadratic transform, so the former converges faster as indicated by the above convergence rate analysis.

We provide an intuitive reason for the preceding conclusion.

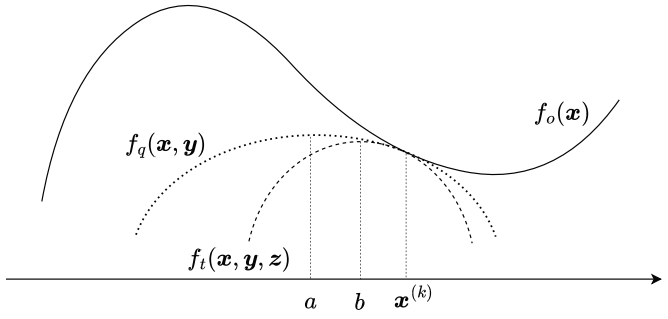


Fig. 12. The basic quadratic transform approximates $f_o(\boldsymbol{x})$ as $f_q(\boldsymbol{x},\boldsymbol{y})$ while the nonhomogeneous quadratic transform approximates $f_o(\boldsymbol{x})$ as $f_t(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$. By the MM procedure, for the current solution $\boldsymbol{x}^{(k)}$, the basic quadratic transform updates it to a, while the nonhomogeneous quadratic transform updates it to b. The basic quadratic transform converges faster in iterations, because its approximation is tighter.

The parameter Λ is proportional to the gap between the surrogate function and the original objective function. More precisely, Λ reflects how well $g(\boldsymbol{x}|\hat{\boldsymbol{x}})$ approximates $f_o(\boldsymbol{x})$ in terms of the second-order profile. In principle, a smaller Λ leads to a tighter approximation. Recall that the basic quadratic transform uses the quadratic transform to construct a lower bound on $f_o(\boldsymbol{x})$, while the nonhomogeneous quadratic transform uses the nonhomogeneous bound [46] to further bound the above lower bound from below, i.e.,

$$f_o(\boldsymbol{x}) \geq f_q(\boldsymbol{x}|\hat{\boldsymbol{x}}) \geq f_t(\boldsymbol{x}|\hat{\boldsymbol{x}}), \tag{83}$$

where $f_q(\boldsymbol{x}|\hat{\boldsymbol{x}})$ is the new objective function used in the basic quadratic transform, while $f_q(\boldsymbol{x}|\hat{\boldsymbol{x}})$ is the new objective function used in the nonhomogeneous quadratic transform. Although both $f_q(\boldsymbol{x}|\hat{\boldsymbol{x}})$ and $f_q(\boldsymbol{x}|\hat{\boldsymbol{x}})$ are surrogate functions of $f_o(\boldsymbol{x})$, the former is closer to $f_o(\boldsymbol{x})$ and gives a tighter approximation as illustrated in Fig. 12, so the basic quadratic transform converges faster. We emphasize that the convergence rate considered here is in terms of the number of iterations. In practice, the nonhomogeneous quadratic transform can run faster than the basic quadratic transform, because it avoids a large matrix inversion in each iteration, even though the nonhomogeneous quadratic transform may require more iterations to converge.

Since Λ measures the gap between the surrogate function and $f_o(\boldsymbol{x})$, the best scenario one can hope for is $\Lambda = 0$. This can be achieved by the *cubically regularized Newton's method* due to Nesterov [53]. As shown in Theorem 4.1.4 in [53], the resulting convergence rate is

$$f_o(\boldsymbol{x}^*) - f_o(\boldsymbol{x}^{(k)}) \leq \frac{LR^3}{2(1+k/3)^2}, \quad \text{for } k \geq 2. \tag{84}$$

However, the cubically regularized Newton's method requires computing the inverse of the Hessian matrix and thus is costly in practice. We now show that the extrapolated quadratic transform can yield almost equally good performance. Recall that the extrapolated quadratic transform accelerates the nonhomogeneous quadratic transform by incorporating Nesterov's extrapolation scheme, so the convergence rate analysis for the momentum method in [53] carries over to the extrapolated quadratic transform method. Suppose that the gradient of $f_o(\boldsymbol{x})$ is C-Lipschitz continuous. Then under certain conditions [53] the extrapolated quadratic transform yields

$$f(\boldsymbol{x}^*) - f(\boldsymbol{x}^{(k)}) \leq \frac{2C\|\boldsymbol{x}^* - \boldsymbol{x}^{(0)}\|_2^2}{(k+1)^2}, \quad \text{for } k \geq 1. \tag{85}$$

As opposed to the basic quadratic transform and nonhomogeneous quadratic transform that guarantee an error bound of $O(1/k)$, the extrapolated quadratic transform provides an improved error bound of $O(1/k^2)$. Moreover, the extrapolated quadratic transform inherits from the nonhomogeneous quadratic transform the advantage of not requiring large matrix inverse operations.

X. Conclusion

Many problems in signal processing and machine learning naturally lead to optimization with fractional structure. This article provides an up-to-date and comprehensive account of an FP technique termed quadratic transform. As opposed to the classic FP methods (e.g., Dinkelbach's method) that are typically limited to the scalar-valued single-ratio problem, the quadratic transform allows for a much wider scope of FP problems, ranging from the sum-of-functions-of-ratio problem to the matrix-ratio problem. Although the quadratic transform is the main focus, this article also covers other related FP techniques, including the Lagrangian dual transform and the AM-GM inequality transform. Moreover, we explore the extensive connections between the quadratic transform and several well-known optimization methods, including the MM method, the gradient projection method, the WMMSE algorithm, and the fixed-point iteration. We further examine the speed of convergence of the quadratic transform. Aside from the theoretical aspect, this article pays much attention to a variety of applications of FP, e.g., the SVM optimization, unsupervised data clustering, AoI minimization, power control, beamforming, and channel estimation.

Acknowledgments

The authors wish to thank the anonymous reviewers for many helpful suggestions, especially on the rational optimization [4] and Lasserre's hierarchy [40].

The work of Kaiming Shen was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 12426306 and in part by the Guangdong Major Project of Basic and Applied Basic Research under Grant 2023B0303000001. The work of Wei Yu was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada via a Discovery Grant.

References

[1] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Nav. Res. Logistics Quart.*, vol. 9, no. 3, pp. 181–186, Dec. 1962.

[2] S. Schaible, "Parameter-free convex equivalent and dual programs of fractional programming problems," *Zeitschrift für Oper. Res.*, vol. 18, no. 5, pp. 187–196, Oct. 1974.

[3] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 133, no. 7, pp. 492–498, Mar. 1967.

[4] D. Jibetean and E. de Klerk, "Global optimization of rational functions: a semidefinite programming approach," *Math. Program.*, vol. 106, pp. 93–109, 2006.

[5] 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, Mar. 2018.

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

[7] J. P. Crouzeix, J. A. Ferland, and S. Schaible, "An algorithm for generalized fractional programs," *J. Optim. Theory Appl.*, vol. 47, no. 1, pp. 35–49, Sep. 1985.

[8] R. W. Freund and F. Jarre, "Solving the sum-of-ratios problem by an interior-point method," *J. Global Optim.*, vol. 19, no. 1, pp. 83–102, 2001.

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

[10] Y. Chen, L. Zhao, and K. Shen, "Mixed max-and-min fractional programming for wireless networks," *IEEE Trans. Signal Process.*, vol. 72, pp. 337–351, Jan. 2024.

[11] 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, Mar. 2018.

[12] 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, Oct. 2019.

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

[14] K. Shen, Z. Zhao, Y. Chen, Z. Zhang, and H. V. Cheng, "Accelerating quadratic transform and WMMSE," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 11, pp. 3110–3124, Nov. 2024.

[15] Z.-A. Liang, H.-X. Huang, and P. M. Pardalos, "Efficiency conditions and duality for a class of multiobjective fractional programming problems," *J. Global Optim.*, vol. 27, pp. 447–471, Dec. 2003.

[16] K. Mathur and M. C. Puri, "On bilevel fractional programming," *Optim.*, vol. 35, no. 3, pp. 215–226, May 1995.

[17] M. Jaberipour and E. Khorram, "Solving the sum-of-ratios problems by a harmony search algorithm," *J. Comput. Appl. Math.*, vol. 234, no. 3, pp. 733–742, Jun. 2010.

[18] H. Arsham and A. B. Kahn, "A complete algorithm for linear fractional programs," *Comput. Math. Appl.*, vol. 20, no. 7, pp. 11–23, 1990.

[19] S. Chakraborty, O. T. Demir, E. Björnson, and P. Giselsson, "Efficient downlink power allocation algorithms for cell-free massive MIMO systems," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 168–186, Jan. 2021.

[20] P. Gu, R. Li, and R. Tafazolli, "Dynamic cooperative spectrum sharing in a multi-beam LEO-GEO co-existing satellite system," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1170–1182, Feb. 2022.

[21] 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, May 2020.

[22] J. Zhang, X. Hu, and C. Zhong, "Phase calibration for intelligent reflecting surfaces assisted millimeter wave communications," *IEEE Trans. Signal Process.*, vol. 70, pp. 1026–1040, Feb. 2022.

[23] N. Su, F. Liu, Z. Wei, Y.-F. Liu, and C. Masouros, "Secure dual-functional radar-communication transmission: Exploiting interference for resilience against target eavesdropping," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7238–7252, Sep. 2022.

[24] J. Yang, G. Cui, X. Yu, and L. Kong, "Dual-use signal design for radar and communication via ambiguity function sidelobe control," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 9781–9793, Sep. 2020.

[25] L. Wu and D. P. Palomar, "Collaborative cloud and edge mobile computing in C-RAN systems with minimal end-to-end latency," *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 259–274, Sep. 2019.

[26] J. Lei and Q. Liu, "Fractional optimization with the learnable prior for electrical capacitance tomography," *IEEE Trans. Comp. Imag.*, vol. 10, pp. 304–317, Feb. 2024.

[27] A. Beck, A. Ben-Tal, and M. Teboulle, "Finding a global optimal solution for a quadratically constrained fractional quadratic problem with applications to the regularized total least squares," *SIAM J. Matrix Anal. Appl.*, vol. 28, no. 2, pp. 425–445, 2006.

[28] 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 Inf. Process. Netw.*, vol. 7, pp. 259–274, Apr. 2021.

[29] Y. He, J. Ren, G. Yu, and J. Yuan, "Importance-aware data selection and resource allocation in federated edge learning system," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13 593–13 605, Nov. 2020.

[30] B. Pirouz and M. Gaudioso, "New mixed integer fractional programming problem and some multi-objective models for sparse optimization," *Soft Comput.*, vol. 27, pp. 15 893–15 904, Jul. 2023.

[31] J. Zhao, L. Qian, and W.-H. Yu, "Human-centric resource allocation in the metaverse over wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 3, pp. 514–537, Mar. 2024.

[32] S. Dias, P. Brito, and P. Amaral, "Discriminant analysis of distributional data via fractional programming," *Eur. J. of Oper. Res.*, vol. 294, pp. 206–218, Jan. 2021.

[33] X. Chen, Z. Xiao, F. Nie, and J. Z. Huang, "FINC: An efficient and effective optimization method for normalized cut," *IEEE Trans. Pattern Anal. Mach. Intell.*, Feb. 2022.

[34] Y. Kawahara, K. Nagano, and Y. Okamoto, "Submodular fractional programming for balanced clustering," *Pattern Recognit. Lett.*, vol. 32, pp. 235–243, 2011.

[35] C. Zhong, H. Yang, and X. Yuan, "Over-the-air federated multi-task learning over MIMO multiple access channels," *IEEE Trans. Wireless Commun.*, vol. 22, no. 6, pp. 3853–3868, Jun. 2023.

[36] 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, Nov. 2020.

[37] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.

[38] F. Bugarin, D. Henrion, and J.-B. Lasserre, "Minimizing the sum of many rational functions," *Math. Program. Comput.*, vol. 8, no. 1, pp. 83–111, Aug. 2015.

[39] J. Nie, "Optimality conditions and finite convergence of Lasserre's hierarchy," *Math. Program.*, vol. 146, pp. 97–121, 2013.

[40] J.-B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM J. Optim.*, vol. 11, no. 3, pp. 796–817, Jan. 2001.

[41] A. Gharanjik, M. Soltanalian, M. R. B. Shankar, and B. Ottersten, "Grab-n-pull: A max-min fractional quadratic programming framework with applications in signal and information processing," *Signal Process.*, vol. 160, Feb. 2019.

[42] T. Kuno, "A branch-and-bound algorithm for maximizing the sum of several linear ratios," *J. Global Optim.*, vol. 22, pp. 155–174, 2002.

[43] S. He, Y. Huang, L. Yang, and B. Ottersten, "Coordinated multicell multiuser precoding for maximizing weighted sum energy efficiency," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 741–751, Feb. 2014.

[44] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 1–14, Jan. 2015.

[45] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.

[46] 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, Aug. 2016.

[47] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.

[48] S. K. Kaul and R. D. Yates, "Age of information: Updates with priority," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 2644–2648.

[49] J. Papandriopoulos and J. S. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009.

[50] Y. Chen, B. Huang, L. Zhao, and K. Shen, "Multidimensional fractional programming for normalized cuts," in *Conf. Neural Inf. Process. Syst. (NeurIPS)*, Dec. 2024.

[51] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Feb. 2000.

[52] F. Nie, J. Lu, D. Wu, R. Wang, and X. Li, "A novel normalized-cut solver with nearest neighbor hierarchical initialization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 1, pp. 659–666, May 2024.

[53] Y. Nesterov, *Lectures on Convex Optimization (Second Edition)*. Springer, 2018.

[54] L. T. K. Hien, D. N. Phan, and N. Gillis, "An inertial block majorization minimization framework for nonsmooth nonconvex optimization," *J. Mach. Learn. Res.*, vol. 24, no. 18, pp. 1–41, Jan. 2023.

[55] S. S. Christensen, R. Argawal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BS beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 1–7, Dec. 2008.
[56] 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, Sep. 2011.
[57] L. Venturino, N. Prasad, and X. Wang, "Coordinated scheduling and power allocation in downlink multicell OFDMA networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 2835–2848, Jul. 2012.
[58] H. Dahrouj, W. Yu, and T. Tang, "Power spectrum optimization for interference mitigation via iterative function evaluation," *EURASIP J. Wireless Commun. Netw.*, Aug. 2012.
[59] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.
[60] D. P. Bertsekas, *Nonlinear Programming (Third Edition)*. Athena Scientific, 2016.

Kaiming Shen received the B.Eng. degree in information security and the B.Sc. degree in mathematics from Shanghai Jiao Tong University, China in 2011, and the Ph.D. degree in electrical and computer engineering from the University of Toronto in 2020. He has been with the School of Science and Engineering at The Chinese University of Hong Kong (Shenzhen), China as a tenure-track assistant professor since 2020. He received the IEEE Signal Processing Society Young Author Best Paper Award in 2021, and the Frontiers of Science Award in 2024. He currently serves as an Editor for IEEE Transactions on Wireless Communications.

Wei Yu received the B.A.Sc. degree in computer engineering and mathematics from the University of Waterloo, Canada, and the Ph.D. degree in electrical engineering from Stanford University, USA. He is currently a Professor in the Electrical and Computer Engineering Department at the University of Toronto, where he holds a Canada Research Chair (Tier 1) in Information Theory and Wireless Communications. He served as the Chair of the Signal Processing for Communications and Networking Technical Committee of the IEEE Signal Processing Society in 2017-2018, and the President of the IEEE Information Theory Society in 2021. Prof. Wei Yu is a Clarivate Highly Cited Researcher. He is a Fellow of IEEE.