Surrogate Optimization for p-Norms*

Yasushi Kawase[1] and Kazuhisa Makino[2]

1 **School of Engineering, Tokyo Institute of Technology, Japan**
`kawase.y.ab@m.titech.ac.jp`
2 **Research Institute for Mathematical Sciences, Kyoto University, Japan**
`makino@kurims.kyoto-u.ac.jp`

Abstract

In this paper, we study the effect of surrogate objective functions in optimization problems. We introduce *surrogate ratio* as a measure of such effect, where the surrogate ratio is the ratio between the optimal values of the original and surrogate objective functions.

We prove that the surrogate ratio is at most $\mu^{|1/p-1/q|}$ when the objective functions are p- and q-norms, and the feasible region is a μ-dimensional space (i.e., a subspace of $\mathbb{R}^\mu$), a μ-intersection of matroids, or a μ-extendible system. We also show that this is the best possible bound. In addition, for μ-systems, we demonstrate that the ratio becomes $\mu^{1/p}$ when $p < q$ and unbounded if $p > q$. Here, a μ-system is an independence system such that for any subset of ground set the ratio of the cardinality of the largest to the smallest maximal independent subset of it is at most μ. We further extend our results to the surrogate ratios for approximate solutions.

1998 ACM Subject Classification G.2.0 General

Keywords and phrases surrogate optimization, matroid, extendible system, p-norm

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2016.41

1 Introduction

When we model real-world problems as mathematical optimization problems, we often face some difficulties choosing appropriate objective functions for the problems. This, for instance, follows from ambiguity and computational difficulty of real objective functions. We demonstrate such examples below. In order to overcome such difficulties, it is natural to make use of surrogate objective functions.

Ambiguity: The first example is for ambiguous objective functions. For instance, consider car navigation system to find a fastest route from an origin to a destination in a road network. It usually solves the shortest path problem with estimated transit time for each road. This is simply because we do not know the actual transit time which depends on traffic congestion. Thus, estimated objective functions are used as surrogate ones.

Computational difficulty: Second, we sometimes approximate the objective function to reduce computational cost. For example, consider the following sparse approximation problem: we are given a vector $b \in \mathbb{R}^m$ and a matrix $A \in \mathbb{R}^{m\times n}$ ($n \gg m$), and we are asked to find a vector x to minimize $\|x\|_0$ ($= |\operatorname{supp}(x)|$) subject to $Ax = b$. Unfortunately, the problem is computationally intractable (NP-hard) [19], and it is often replaced to minimizing $\|x\|_1$ ($= \sum_i^n |x_i|$) or $\|x\|_2$ ($= \sum_i^n x_i^2$) (see, e.g., [20]). The resulting problem can be regarded as a linear or convex programming problem, and thus, we can efficiently solve the surrogate optimization problem.

* The first author is supported by JSPS KAKENHI Grant Number 26887014 and 16K16005. The second author is supported by JSPS KAKENHI Grant.

Multi-objective optimization: Consider a scenario in which we have multiple objectives. In this case, we sometimes use their weighted sum as a surrogate objective function. For instance, in a mean-variance portfolio optimization: given an expected return vector $\bar{p} \in \mathbb{R}^n$ and a variance-covariance matrix $V \in \mathbb{R}^{n \times n}$, we are asked to find a weight vector $x \in \mathbb{R}^n$ ($\sum_{i=1}^n x_i = 1$, $x_i \geq 0$ $\forall i = 1, \ldots, n$) to maximize the expected return $\bar{p}^\top x$ and minimize the risk $x^\top V x$. A standard way to obtain a solution for the problem is to maximize $\bar{p}^\top x - \lambda x^\top V x$, where λ is called a risk-aversion coefficient [4].

Fairness-efficiency trade-off: The next example follows from the trade-off between efficiency and fairness. Consider the following facility location problem: given a set of demand points $D \subseteq V$ in a metric space (V, d), we are asked to select k facilities $F \subseteq V$ to open while minimizing $\sqrt[p]{\sum_{i \in D} (\min_{j \in F} d(i,j))^p}$ for $p \geq 1$. We can see that the case of $p = 1$ (i.e., k-median problem) is efficient and the other extreme case $p = +\infty$ (i.e., k-center problem) is fair. Furthermore, an optimal solution for each p can be regarded as a solution which balances the efficiency and the fairness (see discussion in [7]). Golovin et al. [7] have suggested that the optimal solution for sufficiently large p is good for this trade-off. Namely, sufficient large p provides a good surrogate objective function.

Potential games: The last example is in *potential games*. A game is said to be a potential game if the incentive of all players to change their strategy can be expressed by using a single global function called *potential function* [18]. Potential games always admit a pure Nash equilibrium and, in particular, any minimizer of the potential function is a pure Nash equilibrium. Thus, the potential minimizers are recognized as important solution concept in potential games. On the other hand, the efficiency of a solution is measured by the *social cost*, which is, for example, the sum or maximum of players' cost. For instance, consider the following load balancing game [21]. There are n users N and m identical machines M. Each user $i \in N$ has a job with weight w_i and chooses a machine to place the job. A combination of choices yields an assignment $A : N \to M$. The load of machine $j \in M$ under assignment A is defined as $l_j(A) = \sum_{i \in N : A(i) = j} w_i$. The cost of user $i \in N$ corresponds to the load on machine $A(i)$, i.e., $l_{A(i)}(A)$. Then, a potential for the problem is $\|l(A)\|_2^2$ $(= \sum_{j \in M} l_j(A)^2)$ and a social cost is (usually) the makespan $\|l(A)\|_\infty$ $(= \max_{j \in M} l_j(A))$.

In this paper, we quantify the effect of surrogate objective functions by introducing *surrogate ratio*. The surrogate ratio compares the optimal solutions with respect to the original and surrogate objective functions. Our approach is analogous to the worst-case performance analysis in algorithm theory, such as the approximation ratio, the competitive ratio, and the robustness factor. As the first step in analyzing the surrogate ratio, this paper focuses on optimization problems of maximizing p-norms. Maximizing p-norms are well studied in many areas as shown above. For n-dimensional real vector $x \in \mathbb{R}^n$ and positive real $p \in \mathbb{R}_+$, p-norm of x is defined by $\|x\|_p = \sqrt[p]{\sum_{i=1}^n |x_i|^p}$. We remark that $\|\cdot\|_p$ for p with $p < 1$ is *not* a norm, but we treat $\|\cdot\|_p$ for any positive real p. We also use the notation that $\|x\|_\infty = \lim_{p \to \infty} \|x\|_p = \max_i |x_i|$ for a real vector $x \in \mathbb{R}^n$.

Our model

We discuss the following four types of surrogate ratios between p- and q-norms for maximization problems with a compact non-empty feasible region $S \subseteq \mathbb{R}^\mu$:

$$\rho(S, p, q) = \max\{\|x\|_p : x \in S\} / \min\{\|x\|_p : x \in \arg\max_{x \in S} \|x\|_q\},$$
$$\eta(S, p, q) = \max\{\|x\|_p : x \in S\} / \max\{\|x\|_p : x \in \arg\max_{x \in S} \|x\|_q\},$$

$$\rho_\alpha(S,p,q) = \max\{\|x\|_p : x \in S\} / \min\{\|x\|_p : x \in \alpha\text{-}\arg\max_{x\in S} \|x\|_q\},$$
$$\eta_\alpha(S,p,q) = \max\{\|x\|_p : x \in S\} / \max\{\|x\|_p : x \in \alpha\text{-}\arg\max_{x\in S} \|x\|_q\},$$

where $\alpha \geq 1$ and

$$\alpha\text{-}\arg\max_{x\in S} f(x) = \{x \in S : f(x) \geq f(x')/\alpha \ (\forall x' \in S)\}.$$

Here, we assume for simplicity that the ratios are 1 if $S = \{0\}$. Then any optimal (α-approximate) solution for q-norm is a $\rho(S,p,q)$-approximate ($\rho_\alpha(S,p,q)$-approximate) solution for p-norm and there exists an optimal (α-approximate) solution for q-norm that is an $\eta(S,p,q)$-approximate ($\eta_\alpha(S,p,q)$-approximate) solution for p-norm. The ratios above are respectively called *the worst surrogate ratio*, *the best surrogate ratio*, *the worst α-approximation surrogate ratio*, and *the best α-approximation surrogate ratio*. By definitions, we have

$$1 \leq \eta_\alpha(S,p,q) \leq \eta(S,p,q) \leq \rho(S,p,q) \leq \rho_\alpha(S,p,q)$$

for any S, $p,q > 0$, and $\alpha \geq 1$. Moreover, $\rho(S,p,q) = \rho_1(S,p,q)$ and $\eta(S,p,q) = \eta_1(S,p,q)$ hold. We remark that the surrogate ratios between general two functions can be defined in the same way.

In this paper, we deal with *maximum weight independent set problems*, which are fundamental in combinatorial optimization and contain a number of important problems such as maximum weight stable set problem, maximum weight matching problem, and knapsack problem (see, e.g., [15]). An independence system is a set system $(E,\mathcal{F})$, i.e., E is a finite set and $\mathcal{F}$ is a family of subsets of E, with the following two properties: (I1) $\emptyset \in \mathcal{F}$ and (I2) $Y \subseteq X \in \mathcal{F}$ implies $Y \in \mathcal{F}$. Given an independence system $(E,\mathcal{F})$, a subset F of E is called *independent set* if F belongs to $\mathcal{F}$, and an (inclusion-wise) maximal independent set is called a *base*. For an independence system $(E,\mathcal{F})$ and non-negative weight $w(e)$ for $e \in E$, the maximization problem with p-norm is defined as $\max\{w^p(X) : X \in \mathcal{F}\}$ where we define

$$w^p(X) = \sqrt[p]{\sum\nolimits_{e\in X} w(e)^p}.$$

An independence system $(E,\mathcal{F})$ is called *matroid* if $X, Y \in \mathcal{F}$, $|X| < |Y|$ implies the existence of $v \in Y \setminus X$ such that $X \cup \{v\} \in \mathcal{F}$, and *$\mu$-intersection* if it is an intersection of μ matroids defined over E. As extensions of μ-intersection, we consider μ-extendible systems and μ-systems. An independence system $(E,\mathcal{F})$ is called *μ-extendible*[1] if

$$\forall X, Y \in \mathcal{F},\ \forall e \in Y \setminus X,\ \exists Z \subseteq X \setminus Y \text{ such that } |Z| \leq \mu,\ X \cup \{e\} \setminus Z \in \mathcal{F},$$

and *μ-system* if for all $S \subseteq E$ the ratio of the cardinality of the largest to the smallest maximal independent subset of S is at most μ. For simplicity, we assume that $\mu \geq 1$ in this paper (although we can take $\mu = 0$ for $(E, 2^E)$). It is known that classes of these systems have the following relationships [17]:

$$\mu\text{-intersection} \ \subseteq \ \mu\text{-extendible} \ \subseteq \ \mu\text{-system}.$$

For an independence system $(E,\mathcal{F})$ with a weight $w : E \to \mathbb{R}_+$, we denote the best surrogate ratio as $\rho(E,\mathcal{F},w;p,q)$. We also define $\eta(E,\mathcal{F},w;p,q)$, $\rho_\alpha(E,\mathcal{F},w;p,q)$, and $\eta_\alpha(E,\mathcal{F},w;p,q)$ similarly.

[1] Kakimura and Makino [11] called this system *μ-exchangeable*.

Table 1 Summary of the surrogate ratios for maximization problems.

	μ-dimensional space	μ-intersection	μ-extendible	μ-system
$\rho,\ \eta$	$\mu^{\lvert 1/p-1/q \rvert}$ [Thms. 2, 3]	$\mu^{\lvert 1/p-1/q \rvert}$ [Thm. 6]	$\mu^{\lvert 1/p-1/q \rvert}$ [Thm. 6]	$\begin{cases} \mu^{1/p} & (p < q), \\ \infty & \left(\substack{p>q \\ \mu>1}\right), \\ 1 & (\text{otherwise}) \end{cases}$ [Thm. 13]
η_α $(\alpha > 1)$	$\max\left\{1, \frac{\mu^{\lvert 1/p-1/q \rvert}}{\alpha}\right\}$ [Thm. 3]	$\max\left\{1, \frac{\mu^{\lvert 1/p-1/q \rvert}}{\alpha}\right\}$ [Thm. 3]	$\max\left\{1, \frac{\mu^{\lvert 1/p-1/q \rvert}}{\alpha}\right\}$ [Thm. 3]	$\begin{cases} \mu^{1/p} & (p < q), \\ \infty & \left(\substack{p>q \\ \mu^{1/q}>\alpha}\right), \\ 1 & (\text{otherwise}) \end{cases}$ [Thm. 13]
ρ_α $(\alpha > 1)$	$\alpha \cdot \mu^{\lvert 1/p-1/q \rvert}$ [Thm. 2]	$\begin{cases} \infty & (p \neq q) \\ \alpha & (p = q) \end{cases}$ [Thm. 9]	$\begin{cases} \infty & (p \neq q) \\ \alpha & (p = q) \end{cases}$ [Thm. 9]	$\begin{cases} \infty & (p \neq q) \\ \alpha & (p = q) \end{cases}$ [Thm. 9]

Our results

In this paper, we analyze the surrogate ratios. For a μ-dimensional compact feasible region $S \subseteq \mathbb{R}^\mu$, we show that the best and the worst surrogate ratios are both $\mu^{|1/p-1/q|}$. Analogously, we prove that the best and worst surrogate ratios for a μ-intersection of matroids and a μ-extendible system are $\mu^{|1/p-1/q|}$. On the other hand, for a μ-system with $\mu > 1$, we cannot bound the surrogate ratios by $\mu^{|1/p-1/q|}$. The ratios become $\mu^{1/p}$ if $p < q$, and unbounded if $p > q$. Note that, optimality of a matroid (when $\mu = 1$) is independent of p, since the greedy algorithm, which always produces an optimal solution, does not need the values of weights but the ordering of them. Thus, the surrogate ratios are 1 in this case. Moreover, for any independence system, the greedy solution[2] coincides with an optimal solution for a p-norm with a sufficiently large p. Our result for μ-systems implies that the greedy solution is a μ-approximate solution by choosing $p = 1$ and q as a sufficiently large number, that is also shown by Jenkyns [10] and Korte and Hausmann [14]. The best α-approximation surrogate ratio is basically α times smaller than the previous one, i.e. the case $\alpha = 1$. On the contrary, the worst α-approximation surrogate ratio goes to infinity except for the μ-dimensional compact case.

Our results are summarized as Table 1.

Related work

The surrogate ratio is not just an analogy of the ratios shown below, but also closely related to them.

For a multi-objective or robust optimization problem, a natural measure of goodness of a solution is the ratio between the value of the solution and the optimal one for the worst objective function. The ratio is called *robustness factor* [9] (also studied under the name of *simultaneous approximation ratio* [6] or *global approximation ratio* [16]). To be more precise, assume that we want to maximize $f_1, \dots, f_n$ under the constraint $x \in S$. Then a solution $x \in S$ is β-robust if $f_i(x) \geq f_i(y)/\beta$ holds for all $y \in S$ and $i = 1, \dots, n$. Thus, we can obtain a β-robust solution by maximizing g, if the surrogate ratio is at most β for each f_i and g.

[2] To be precise, "greedy solution" may not be unique when there exist ties in the weights. However, we can perturb the weights slightly to break the ties with arbitrarily small changes in the values of p-norms.

Azar et al. [3] introduced a concept of an all-norm ρ-approximation algorithm, which supplies one solution that guarantees ρ-approximation to p-norms simultaneously. They gave a 2-approximation polynomial time algorithm for the p-norm load balancing problem (or rather, the problem of restricted assignment model).

Hassin and Rubinstein [9] studied a robustness version of maximum weight independent set problem when the objective functions are the sum of the k largest weights of selected elements for all positive integer k. They showed that, when the family of independent set is that of matchings in a graph, the optimal solution for p-norm ($p \geq 1$) is $\min\{2^{(1/p)-1}, 2^{-1/p}\}$-robust. This implies the existence of $\sqrt{2}$-robust solution for any graph by choosing $p = 2$. They also proved that $\sqrt{2}$-robust is the best possible. Fujita et al. [5] extended the result to the matroid intersection case. Kakimura and Makino [11] further extended the result to the μ-extendible system and showed that the optimal solution with respect to p-norm is $\min\{\mu^{(1/p)-1}, \mu^{-1/p}\}$-robust. We remark that this result does not imply our results.

For a potential game, consider a surrogate ratio of a social cost and a potential function. Then the surrogate ratio can be used to quantify the inefficiency of selfish behavior in the game. In fact, the ratio is studied under the name of the *inefficiency ratio of stable equilibria* [2] or the *potential-optimal price of anarchy* [13]. Note that, in algorithmic game theory, one of the most famous measures to quantify the inefficiency of a game is the *price of stability*, which is the ratio of the social cost at the best equilibrium to the minimum social cost possible [1]. An upper bound on the price of stability is often calculated by using the surrogate ratio. This bounding technique is called *potential function method* [1]. Moreover, we can see that the price of stability is a surrogate ratio where we do not replace the objective function but the feasible region is restricted from all the possible strategy profiles to the set of equilibria.

2 Surrogate ratios for μ-dimensional space

In this section, we study the surrogate ratios for μ-dimensional space. The following proposition plays an important role to obtain upper bounds on the surrogate ratios.

▶ **Proposition 1** (Norm Inequalities (see, e.g., [8])). *For any n-dimensional vector $x \in \mathbb{R}^n$ and $0 < p \leq q \leq \infty$, it holds that $\|x\|_q \leq \|x\|_p \leq n^{\frac{1}{p}-\frac{1}{q}}\|x\|_q$.*

We first evaluate the worst (α-approximation) surrogate ratio.

▶ **Theorem 2.** *For any $0 < p, q \leq \infty$ and $\alpha \geq 1$, we have*

$$\max_{S \subseteq \mathbb{R}^\mu:\substack{\textit{non-empty}\\ \textit{compact}}} \frac{\max\{\|x\|_p : x \in S\}}{\min\{\|x\|_p : x \in \alpha\text{-}\arg\max_{x\in S}\|x\|_q\}} = \alpha \cdot \mu^{|1/q-1/p|}. \tag{1}$$

Proof. We first claim that the left hand side of (1) is upper bounded by $\alpha \cdot \mu^{|1/q-1/p|}$. Let $M = \max\{\|x\|_q : x \in S\}$. Then we have

$$\frac{\max\{\|x\|_p : x \in S\}}{\min\{\|x\|_p : x \in \alpha\text{-}\arg\max_{x\in S}\|x\|_q\}} \leq \frac{\max\{\|x\|_p : \|x\|_q \leq M\}}{\min\{\|x\|_p : \|x\|_q \geq M/\alpha\}}.$$

By Proposition 1, if $p \leq q$, we have $\max\{\|x\|_p : \|x\|_q \leq M\} \leq \mu^{1/p-1/q} \cdot M$ and $\min\{\|x\|_p : \|x\|_q \geq M/\alpha\} \geq M/\alpha$. Otherwise (i.e., $p > q$), we have $\max\{\|x\|_p : \|x\|_q \leq M\} \leq M$ and $\min\{\|x\|_p : \|x\|_q \geq M/\alpha\} \geq \mu^{1/p-1/q} \cdot M/\alpha$. Therefore, we obtain $\max\{\|x\|_p : \|x\|_q \leq M\}/\min\{\|x\|_p : \|x\|_q \geq M/\alpha\} = \alpha \cdot \mu^{|1/q-1/p|}$.

Next, we show that there exists a μ-dimensional compact set S that attains the maximum in (1). Let $A_\gamma = \{a_\gamma, \mathbf{1}\}$ where $a_\gamma = (\gamma \cdot \mu^{1/q}, 0, \ldots, 0)^\top \in \mathbb{R}^\mu$ and $\mathbf{1} = (1, 1, \ldots, 1)^\top \in \mathbb{R}^\mu$.

Then, we can observe that the worst α-approximation surrogate ratio of A_α or $A_{1/\alpha}$ is $\alpha \cdot \mu^{|1/q-1/p|}$. ◀

This theorem yields that any α-approximate solution for q-norm is $\alpha \cdot \mu^{|1/q-1/p|}$-approximate solution for p-norm.

Next, we evaluate the best (α-approximation) surrogate ratio.

▶ **Theorem 3.** *For any $0 < p, q \leq \infty$ and $\alpha \geq 1$, we have*

$$\sup_{S \subseteq \mathbb{R}^\mu : \substack{\textit{non-empty} \\ \textit{compact}}} \frac{\max\{\|x\|_p : x \in S\}}{\max\{\|x\|_p : x \in \alpha\text{-}\arg\max_{x \in S} \|x\|_q\}} = \max\{1, \mu^{|1/q-1/p|}/\alpha\}. \tag{2}$$

Proof. We first claim that $\max\{\|x\|_p : x \in S\} / \max\{\|x\|_p : x \in \alpha\text{-}\arg\max_{x \in S} \|x\|_q\} \leq \max\{1, \mu^{|1/q-1/p|}/\alpha\}$ holds for any non-empty, compact set $S \subseteq \mathbb{R}^\mu$. Let $x_p \in \arg\max_{x \in S} \|x\|_p$. If $\alpha \geq \mu^{|1/q-1/p|}$, then $x_p \in \alpha\text{-}\arg\max_{x \in S} \|x\|_q$ by Theorem 2 (with $\alpha = 1$). Thus, $\max\{\|x\|_p : x \in S\} / \max\{\|x\|_p : x \in \alpha\text{-}\arg\max_{x \in S} \|x\|_q\} = 1$. Otherwise, i.e., $\mu^{|1/q-1/p|} > \alpha \geq 1$, let r satisfies $\alpha = \mu^{|1/q-1/r|}$ and $\min\{p, q\} \leq r \leq \max\{p, q\}$, and let $x_r \in \arg\max_{x \in S} \|x\|_r$. Then $x_r \in \alpha\text{-}\arg\max_{x \in S} \|x\|_q$ by Theorem 2 and hence

$$\begin{aligned}
\frac{\max\{\|x\|_p : x \in S\}}{\max\{\|x\|_p : x \in \alpha\text{-}\arg\max_{x \in S} \|x\|_q\}} &\leq \frac{\max\{\|x\|_p : x \in S\}}{\|x_r\|_p} \\
&\leq \frac{\max\{\|x\|_p : x \in S\}}{\min\{\|x\|_p : x \in \arg\max_{x \in S} \|x\|_r\}} \\
&\leq \mu^{|1/r-1/p|} = \frac{\mu^{|1/q-1/p|}}{\alpha}
\end{aligned}$$

where the last inequality holds by Theorem 2.

Conversely, the best α-approximation surrogate ratio of $A_{\alpha+\varepsilon}$ or $A_{1/(\alpha+\varepsilon)}$ in the proof of Theorem 2 converges to $\max\{1, \mu^{|1/q-1/p|}/\alpha\}$ as ε goes to $+0$. Thus, we obtain the theorem. ◀

This theorem implies that there exists $x \in (\alpha\text{-}\arg\max_{x \in S} \|x\|_p) \cap (\beta\text{-}\arg\max_{x \in S} \|x\|_q)$ if $\alpha\beta \geq \mu^{|1/p-1/q|}$.

3 Independence systems

In this section, we study some properties of independence systems.

For an independence system $(E, \mathcal{F})$ and $A \subseteq E$, define $\mathcal{F}|A = \{X : A \supseteq X \in \mathcal{F}\}$. Then $(A, \mathcal{F}|A)$ is called the *restriction* of $(E, \mathcal{F})$ to A. Also, define for $(E, \mathcal{F})$ and $A \subseteq E$, $\mathcal{F} \setminus A = \{X \setminus A : X \in \mathcal{F}\}$ and $\mathcal{F}/A = \{X \setminus A : A \subseteq X \in \mathcal{F}\}$. Then $(E \setminus A, \mathcal{F} \setminus A)$ and $(E \setminus A, \mathcal{F}/A)$ are called the *deletion* and the *contraction* of $(E, \mathcal{F})$ by A, respectively. If an independence system $(E, \mathcal{F})$ is μ-extendible, then $(A, \mathcal{F}|A)$, $(A, \mathcal{F} \setminus A)$, and $(E \setminus A, \mathcal{F}/A)$ are also μ-extendible [11].

Since a μ-extendible system is a μ-system, we have the following proposition.

▶ **Proposition 4.** *If $(E, \mathcal{F})$ is a μ-extendible system, then we have $|X| \leq \mu \cdot |Y|$ for any bases X, Y of $(E, \mathcal{F})$.*

Next, we see that the supremum of the worst surrogate ratio coincides with that of the best surrogate ratio for any independence system.

▶ **Lemma 5.** *For any independence system* $(E, \mathcal{F})$ *and* $p, q > 0$*, we have*

$$\sup_{w:E\to\mathbb{R}_+} \rho(E,\mathcal{F},w;p,q) = \sup_{w':E\to\mathbb{R}_+} \eta(E,\mathcal{F},w';p,q). \tag{3}$$

4 Surrogate ratios for μ-intersection and μ-extendible systems

In this section, we provide the maximum value of the surrogate ratio for μ-intersection and μ-extendible system. By Lemma 5, we only need to consider the worst one. We remind that we use the notation $\rho(E,\mathcal{F},w;p,q) = \frac{\max\{w^p(X):X\in\mathcal{F}\}}{\min\{w^p(X):X\in\arg\max_{X\in\mathcal{F}} w^q(X)\}}$ for the worst surrogate ratio.

We show that the tight bound of the surrogate ratio is $\mu^{|1/q-1/p|}$.

▶ **Theorem 6.** *For any* $0 < p \le \infty$ *and* $0 < q < \infty$*, we have*

$$\max_{\substack{(E,\mathcal{F}):\mu\text{-intersection}\\ w:E\to\mathbb{R}_+}} \rho(E,\mathcal{F},w;p,q) = \max_{\substack{(E,\mathcal{F}):\mu\text{-extendible}\\ w:E\to\mathbb{R}_+}} \rho(E,\mathcal{F},w;p,q) = \mu^{|1/q-1/p|}.$$

We first provide the lower bound.

▶ **Lemma 7.** *For any* $0 < p \le \infty$*,* $0 < q < \infty$*, and* μ (≥ 1)*, there exists a* μ*-intersection of matroids* $(E, \mathcal{F})$ *and a weight* $w : E \to \mathbb{R}_+$ *such that*

$$\rho(E,\mathcal{F},w;p,q) = \mu^{|1/q-1/p|}.$$

Proof. Let $\mathcal{F} = \{X : X = \{e_0\}$ or $X \subseteq L\}$ for $L = \{e_1,\dots,e_\mu\}$. Here, $(E,\mathcal{F})$ is a μ-intersection of matroids. In fact, $\mathcal{F} = \bigcap_{i=1}^{\mu} \mathcal{F}_i$ holds for partition matroids $\mathcal{F}_i = \{X \subseteq \{e_0,e_1,\dots,e_\mu\} : |X\cap\{e_0,e_i\}| \le 1\}$. Let $w(e_0) = \mu^{1/q}$, $w(e_1) = w(e_2) = \cdots = w(e_l) = 1$. Then $w^p(\{e_0\}) = w^q(\{e_0\}) = \mu^{1/q}$ and $w^p(L) = \mu^{1/p}$, $w^q(L) = \mu^{1/q}$. Thus, the surrogate ratio is

$$\frac{\max\{w^p(X) : X \in \mathcal{F}\}}{\min\{w^p(X) : X \in \arg\max_{X\in\mathcal{F}} w^q(X)\}} = \frac{\max\{\mu^{1/p},\mu^{1/q}\}}{\min\{\mu^{1/p},\mu^{1/q}\}} = \mu^{|1/q-1/p|},$$

which proves the lemma. ◀

We next present the upper bound.

▶ **Lemma 8.** *For any* $0 < p \le \infty$ *and* $0 < q < \infty$*, any* μ*-extendible independence system* $(E,\mathcal{F})$ $(\mu \ge 1)$ *and any weight* $w : E \to \mathbb{R}_+$*, we have*

$$\rho(E,\mathcal{F},w;p,q) \le \mu^{|1/q-1/p|}.$$

Proof. We assume that $p \ne q$ since the claim is obvious for the case $p = q$. We show $\rho(E,\mathcal{F},w;p,q) \le \mu^{|1/q-1/p|}$ for any μ-extendible system $(E,\mathcal{F})$ and any weight $w : E \to \mathbb{R}_+$ by contradiction. We only prove the case $p > q$ because the case $p < q$ can be observed in a similar way. Assume that there exists a μ-extendible system $(E,\mathcal{F})$ and a weight $w : E \to \mathbb{R}_+$ such that $\rho(E,\mathcal{F},w;p,q) > \mu^{|1/q-1/p|}$. We choose such $(E,\mathcal{F})$ so that $|E|$ is as small as possible and w so that $\rho(E,\mathcal{F},w;p,q)$ is maximal for $(E,\mathcal{F})$. Such a w exists since we can pick

$$w \in \arg\max\left\{u^p(X_p) : \substack{u^p(X_q)=1,\ u^q(X_q)\ge u^q(X)\ (\forall X\in\mathcal{F}),\ u_e\ge 0\ (\forall e\in E),\\ X_p,X_q\in\mathcal{F}}\right\},$$

where the objective function is continuous and the feasible region is non-empty and compact. Here, the feasible region is bounded because $u_e \leq u^q(X_q) \leq |X_q|^{|1/q-1/p|} \cdot u^p(X_q) \leq |E|^{|1/q-1/p|}$ for each $e \in E$ and closed because intersection or union of finitely many closed sets is closed. Let

$$X_p \in \operatorname*{arg\,max}_{X \in \mathcal{F}} w^p(X) \quad \text{and} \quad X_q \in \arg\min\{w^p(X) : X \in \operatorname*{arg\,max}_{X \in \mathcal{F}} w^q(X)\}.$$

Without loss of generality, we may assume X_p and X_q are bases of $(E, \mathcal{F})$. We consider the following seven cases.

Case 1. $|X_p| = 1$. Let $X_p = \{e^*\}$. Then $|X_q| \leq \mu$ by Proposition 4. Therefore, we have that the surrogate ratio is at most

$$\frac{w^p(X_p)}{w^p(X_q)} \leq \frac{w(e^*)}{\min\{u^p(X_q) : u^q(X_q) \geq w(e^*)\}} = \frac{w(e^*)}{\mu^{1/p-1/q} \cdot w(e^*)} = \mu^{|1/q-1/p|}$$

where the second equality holds by the norm inequality (Proposition 1) and $|X_q| \leq \mu$.

Case 2. $w(e) = 0$ for some $e \in E$. In this case, $\rho(E, \mathcal{F}, w; p, q) = \rho(E \setminus \{e\}, \mathcal{F} \setminus \{e\}, w; p, q)$ holds since $X_p \setminus \{e\} \in \arg\max_{X \in \mathcal{F} \setminus \{e\}} w^p(X)$ and $X_q \setminus \{e\} \in \arg\min\{w^p(X) : X \in \arg\max_{X \in \mathcal{F} \setminus \{e\}} w^q(X)\}$. This contradicts the minimality of $|E|$.

Case 3. $X_p \cup X_q \subsetneq E$. Let $e \in E \setminus (X_p \cup X_q)$. Then $\rho(E, \mathcal{F}, w; p, q) \leq \rho(E \setminus \{e\}, \mathcal{F} \setminus \{e\}, w; p, q)$ since $X_p \in \operatorname*{arg\,max}_{X \in \mathcal{F} \setminus \{e\}} w^p(X)$ and $X_q \in \arg\min\{w^p(X) : X \in \operatorname*{arg\,max}_{X \in \mathcal{F} \setminus \{e\}} w^q(X)\}$. This contradicts the minimality of $|E|$.

Case 4. $X_p \cap X_q \neq \emptyset$. Let $e \in X_p \cap X_q$. Then $\rho(E, \mathcal{F}, w; p, q) \leq \rho(E \setminus \{e\}, \mathcal{F}/\{e\}, w; p, q)$ since $X_p \setminus \{e\} \in \operatorname*{arg\,max}_{X \in \mathcal{F}/\{e\}} w^p(X)$ and $X_q \setminus \{e\} \in \arg\min\{w^p(X) : X \in \operatorname*{arg\,max}_{X \in \mathcal{F}/\{e\}} w^q(X)\}$. This contradicts the minimality of $|E|$.

Case 5. There exists $X'_p \in \arg\max_{X \in \mathcal{F}} w^p(X)$ such that $X'_p \neq X_p$. We may assume that $w(e) > 0$ for any $e \in E$ by Case 2, $X_p \cup X_q = E$ by Case 3, and $X_p \cap X_q = \emptyset$ by Case 4. Then $X'_p \cup X_q \subsetneq X_p \cup X_q = E$ holds because $X'_p \cup X_q = E$ implies $X'_p \supsetneq X_p$ and $w(X'_p) > w(X_p)$, a contradiction. Thus we have $X'_p \in \arg\max_{X \in \mathcal{F}|(X'_p \cup X_q)} w^p(X)$ and $X_q \in \arg\min\{w^p(X) : X \in \arg\max_{X \in \mathcal{F}|(X'_p \cup X_q)} w^q(X)\}$. This contradicts the minimality $|E|$.

Case 6. There exists $X'_q \in \arg\max_{X \in \mathcal{F}} w^q(X)$ such that $X'_q \notin \{X_p, X_q\}$. We may assume that $w(e) > 0$ for any $e \in E$ by Case 2, and $X_p \cup X_q = E$ and $X_p \cap X_q = \emptyset$ by Cases 3 and 4. Then $w^p(X_p)/w^p(X_q)$ is at most

$$\sqrt[p]{\max\left\{\frac{\sum_{e \in X_p \setminus X'_q} w(e)^p}{\sum_{e \in X_q \cap X'_q} w(e)^p}, \frac{\sum_{e \in X_p \cap X'_q} w(e)^p}{\sum_{e \in X_q \setminus X'_q} w(e)^p}\right\}} = \max\left\{\frac{w^p(X_p \setminus X'_q)}{w^p(X_q \cap X'_q)}, \frac{w^p(X_p \cap X'_q)}{w^p(X_q \setminus X'_q)}\right\}$$

by the mediate inequality. Let $\mathcal{F}_1 = (\mathcal{F}|(X_p \cup X'_q))/(X_p \cap X'_q)$ and $\mathcal{F}_2 = (\mathcal{F}|(X_q \cup X'_q))/(X_q \cap X'_q)$. Then $X_p \setminus X'_q \in \arg\max_{X \in \mathcal{F}_1} w^p(X)$, $X_q \cap X'_q \in \arg\max_{X \in \mathcal{F}_1} w^q(X)$, $X_p \cap X'_q \in \mathcal{F}_2$, and $X_q \setminus X'_q \in \arg\max_{X \in \mathcal{F}_2} w^q(X)$. Thus, we have

$$\frac{w^p(X_p \setminus X'_q)}{w^p(X_q \cap X'_q)} \le \frac{\max_{X \in \mathcal{F}_1} w^p(X)}{\min\{w^p(X) : X \in \arg\max_{X \in \mathcal{F}_1} w^q(X)\}} \le \mu^{|1/q-1/p|},$$
$$\frac{w^p(X_p \cap X'_q)}{w^p(X_q \setminus X'_q)} \le \frac{\max_{X \in \mathcal{F}_2} w^p(X)}{\min\{w^p(X) : X \in \arg\max_{X \in \mathcal{F}_2} w^q(X)\}} \le \mu^{|1/q-1/p|}$$

by the minimality of $|E|$ and hence we have $w^p(X_p)/w^p(X_q) \le \mu^{|1/q-1/p|}$, a contradiction.

Case 7. The other case, i.e., $|X_p| \ge 2$, $w^p(X_p) > w^p(X)$ for any $X \in \mathcal{F} \setminus \{X_p\}$, $w^q(X_q) > w^q(X)$ for any $X \in \mathcal{F} \setminus \{X_p, X_q\}$, $X_p \cap X_q = \emptyset$, $X_p \cup X_q = E$, and $w(e) > 0$ for any $e \in E$. Let $s, t \in X_p$ such that $s \ne t$ and $w(s) \ge w(t)$. For a sufficiently small positive number ε, define

$$\hat{w}_e = \begin{cases} w(e) & (e \in E \setminus \{s,t\}), \\ (w(e)^q + \varepsilon)^{1/q} & (e = s), \\ (w(e)^q - \varepsilon)^{1/q} & (e = t). \end{cases}$$

Recall that $q < \infty$. Then $X_p \in \arg\max_{X \in \mathcal{F}} \hat{w}^p(X_p)$ and $\{X_q\} = \arg\max_{X \in \mathcal{F}} \hat{w}^q(X)$. Here, $\hat{w}^p(X_p) > w^p(X_p)$ and $\hat{w}^p(X_q) = w^p(X_q)$. Thus $\rho(E, \mathcal{F}, w; p, q) = w^p(X_p)/w^p(X_q) < \hat{w}^p(X_p)/\hat{w}^p(X_q) = \rho(E, \mathcal{F}, \hat{w}; p, q)$, which contradicts the maximality of $\rho(E, \mathcal{F}, w; p, q)$. ◀

By Lemmas 7 and 8, we obtain Theorem 6.

5 Worst α-approximation surrogate ratio

In this section, we prove that the worst α-approximation surrogate ratio ($\alpha > 1$)

$$\rho_\alpha(E, \mathcal{F}, w; p, q) = \frac{\max\{w^p(X) : X \in \mathcal{F}\}}{\min\{w^p(X) : X \in \alpha\text{-}\arg\max_{X \in \mathcal{F}} w^q(X)\}}$$

is unbounded even if $(E, \mathcal{F})$ is a free matroid (i.e., $\mathcal{F} = 2^E$), when $p \ne q$.

▶ **Theorem 9.** *For any $\alpha > 1$ and $0 < p \le \infty$, $0 < q < \infty$ ($p \ne q$), there exists a sequence of matroids $(E_k, \mathcal{F}_k)$ and non-negative weights $w_k : E_k \to \mathbb{R}_+$ such that* $\lim_{k\to\infty} \rho_\alpha(E_k, \mathcal{F}_k, w_k; p, q) = \infty$.

Proof. Let $E = \{e_1, \dots, e_k, e_{k+1}\}$ and let $w_k(e_1) = \cdots = w_k(e_k) = 1$, $w_k(e_{k+1}) = \sqrt[q]{(\alpha^q - 1) \cdot k}$. Define $\mathcal{F} = 2^E$. Then $\max_{X \in \mathcal{F}} w_k^p(X) = w_k^p(E) = \sqrt[p]{((\alpha^q-1)\cdot k)^{p/q} + k}$ and $\max_{X \in \mathcal{F}} w_k^q(X) = w_k^q(E) = \alpha \cdot k^{1/q}$. Here, $A = \{e_1, \dots, e_k\}$ is an α-approximate solution for w_k^q since $w_k^q(A) = k^{1/q}$. Thus, the surrogate ratio of α-approximation is at least

$$\frac{w_k^p(E)}{w_k^p(A)} = \frac{\sqrt[p]{((\alpha^q-1)\cdot k)^{p/q} + k}}{k^{1/p}} = \sqrt[p]{(\alpha^q - 1)^{p/q} \cdot k^{p/q-1} + 1} \to \infty \quad (k \to \infty)$$

when $p > q > 0$.

The proof for the case $q > p > 0$ is similar. ◀

We remark that $\rho_\alpha(E, \mathcal{F}, w; p, p) \le \alpha$ holds by the definition. In addition, it holds that $\rho_\alpha(E, \mathcal{F}, w; p, p) = \alpha$ when $E = \{x, y\}$, $\mathcal{F} = 2^E$, and $w(x) = \alpha - 1$, $w(y) = 1$.

6 Best α-approximation surrogate ratio

In this section, we provide the best α-approximation surrogate ratio ($\alpha > 1$)

$$\eta_\alpha(E,\mathcal{F},w;p,q) = \frac{\max\{w^p(X) : X \in \mathcal{F}\}}{\max\{w^p(X) : X \in \alpha\text{-}\arg\max_{X\in\mathcal{F}} w^q(X)\}}$$

for μ-intersection of matroids and μ-extendible systems.

▶ **Theorem 10.** *For any $0 < p \le \infty$, $0 < q < \infty$, and $\alpha \ge 1$, we have*

$$\sup_{\substack{(E,\mathcal{F}):\mu\text{-intersection}\\ w:E\to\mathbb{R}_+}} \eta_\alpha(E,\mathcal{F},w;p,q) = \sup_{\substack{(E,\mathcal{F}):\mu\text{-extendible}\\ w:E\to\mathbb{R}_+}} \eta_\alpha(E,\mathcal{F},w;p,q) = \max\left\{1, \frac{\mu^{|1/q-1/p|}}{\alpha}\right\}.$$

We first provide the lower bound.

▶ **Lemma 11.** *For any $0 < p \le \infty$, $0 < q < \infty$, and $\alpha \ge 1$, integer μ (≥ 1), and $\varepsilon > 0$, there exists a μ-intersection of matroids $(E,\mathcal{F})$ and a weight $w : E \to \mathbb{R}_+$ such that*

$$\frac{\max\{w^p(X) : X \in \mathcal{F}\}}{\max\{w^p(X) : X \in \alpha\text{-}\arg\max_{X\in\mathcal{F}} w^q(X)\}} = \max\{1, \mu^{|1/q-1/p|}/(\alpha+\varepsilon)\}.$$

Proof. Let $\mathcal{F} = \{X : X = \{e_0\} \text{ or } X \subseteq B\}$ for $B = \{e_1,\dots,e_\mu\}$. Here, $(E,\mathcal{F})$ can be viewed as μ-intersection of matroids. In fact, $\mathcal{F} = \bigcap_{i=1}^{\mu} \mathcal{F}_i$ when $\mathcal{F}_i = \{X \subseteq E : |X \cap \{e_0, e_i\}| \le 1\}$. Then the lemma holds, for the case $p < q$, by setting $w(e_0) = (\alpha+\varepsilon)\cdot\mu^{1/q}$ and $w(e_1) = w(e_2) = \cdots = w(e_\mu) = 1$. Also, for the case $p > q$, we can observe the lemma by analyzing the weights $u(e_0) = \mu^{1/q}/(\alpha+\varepsilon)$ and $u(e_1) = u(e_2) = \cdots = u(e_\mu) = 1$. ◀

We next present the upper bound.

▶ **Lemma 12.** *For any $p, q > 0$, $\alpha \ge 1$, a μ-extendible independence system $(E,\mathcal{F})$ and a weight $w : E \to \mathbb{R}_+$, we have*

$$\frac{\max\{w^p(X) : X \in \mathcal{F}\}}{\max\{w^p(X) : X \in \alpha\text{-}\arg\max_{X\in\mathcal{F}} w^q(X)\}} \le \max\{1, \mu^{|1/q-1/p|}/\alpha\}.$$

Proof. Let $X_p \in \arg\max_{X\in\mathcal{F}} w^p(X)$ and $X_q \in \arg\max_{X\in\mathcal{F}} w^q(X)$. If $\alpha \ge \mu^{|1/q-1/p|}$, then $X_p \in \alpha\text{-}\arg\max_{X\in\mathcal{F}} w^q(X)$ by Lemma 8. Thus, $\max\{w^p(X) : X \in \mathcal{F}\}/\max\{w^p(X) : X \in \alpha\text{-}\arg\max_{X\in\mathcal{F}} w^q(X)\} = 1$.

Otherwise, i.e., $\mu^{|1/q-1/p|} > \alpha \ge 1$, let r satisfies $\alpha = \mu^{|1/q-1/r|}$ and $\min\{p,q\} \le r \le \max\{p,q\}$, and let $X_r \in \arg\max_{X\in\mathcal{F}} w^r(X)$. Then $X_r \in \alpha\text{-}\arg\max_{X\in\mathcal{F}} w^q(X)$ by Lemma 8 and hence

$$\begin{aligned}\frac{\max\{w^p(X) : X \in \mathcal{F}\}}{\max\{w^p(X) : X \in \alpha\text{-}\arg\max_{X\in\mathcal{F}} w^q(X)\}} &\le \frac{\max\{w^p(X) : X \in \mathcal{F}\}}{w^p(X_r)}\\ &\le \frac{\max\{w^p(X) : X \in \mathcal{F}\}}{\min\{w^p(X) : X \in \arg\max_{X\in\mathcal{F}} w^r(X)\}}\\ &\le \mu^{|1/r-1/p|} = \frac{\mu^{|1/q-1/p|}}{\alpha}\end{aligned}$$

where the last inequality holds by Lemma 8. ◀

By Lemmas 11 and 12, we get Theorem 10.

7 Surrogate ratios for μ-system

In this section, we show the surrogate ratios for μ-systems. By Theorem 9, the worst α-approximation surrogate ratio ($\alpha > 1$) goes to infinity when $p \neq q$. Hence, we here only analyze the best (approximation) surrogate ratio.

▶ **Theorem 13.** *For any $p, q > 0$ and $\alpha \geq 1$, we have*

$$\sup_{\substack{(E,\mathcal{F}):\mu\text{-system}\\ w:E\to\mathbb{R}_+}} \eta_\alpha(E,\mathcal{F},w;p,q) = \begin{cases} \mu^{1/p} & (p < q),\\ \infty & (p > q,\ \mu^{1/q} > \alpha),\\ 1 & (\textit{otherwise}).\end{cases}$$

We first prove the lower bound.

▶ **Lemma 14.** *For any $p, q > 0$, μ (≥ 1), and $\alpha \geq 1$, there exists a sequence of μ-systems $(E_k, \mathcal{F}_k)$ and weights $w_k : E_k \to \mathbb{R}_+$ ($k = 1, 2, \dots$) such that*

$$\lim_{k\to\infty} \eta_\alpha(E_k,\mathcal{F}_k,w_k;p,q) = \begin{cases} \mu^{1/p} & (p < q),\\ \infty & (p > q,\ \mu^{1/q} > \alpha).\end{cases}$$

Proof. Let $E_k = \{e_1, e_2, \dots, e_{k\cdot\mu}, f\}$, $\mathcal{F}_k = \{F \subseteq E_k : f \notin F \text{ or } |F| \leq k\}$. Then $(E_k, \mathcal{F}_k)$ is a μ-system. Let $X = \{e_1, \dots, e_{k\cdot\mu}\}$ and $Y_\sigma = \{f, e_{\sigma(1)}, \dots, e_{\sigma(k-1)}\}$ $(1 \leq \sigma(1) < \dots < \sigma(k-1) \leq k\cdot\mu)$. We can see the lemma, for the case $p > q > 0$ and $\mu^{1/q} > \alpha \geq 1$, by setting $w_k(e_i) = 1$ $(i = 1, \dots, k\cdot\mu)$, $w_k(f) = \sqrt[q]{k\cdot(\mu/\beta - 1) + 1}$ where β is an arbitrary number such that $\mu > \beta > \alpha^q$. Also we can observe the lemma, for the case $q > p > 0$, by choosing $w_k(e_i) = 1$ $(i = 1, \dots, k\cdot\mu)$ and $w_k(f) = \alpha\sqrt[q]{k\cdot\mu}$. ◀

We next provide the upper bound for the case $p < q$.

▶ **Lemma 15.** *For any $q > p > 0$, μ-system $(E, \mathcal{F})$ ($\mu \geq 1$), and weight $w : E \to \mathbb{R}_+$, we have $\rho(E, \mathcal{F}, w; p, q) \leq \mu^{1/p}$.*

Proof. Let $X_q = \{a_1, \dots, a_k\} \in \arg\min\{w^p(X) : X \in \arg\max_{X\in\mathcal{F}} w^q(X)\}$ and $X_p = \{b_1, \dots, b_l\} \in \arg\max_{X\in\mathcal{F}} w^p(X)$. Without loss of generality, we may assume X_p and X_q are bases. Thus, we have $l \leq \mu\cdot k$ because $(E, \mathcal{F})$ is a μ-system. We additionally assume that $w(a_1) \geq w(a_2) \geq \dots \geq w(a_k)$ and $w(b_1) \geq w(b_2) \geq \dots \geq w(b_l)$. For simplicity, define $w(b_i) = 0$ for $i > l$. Since $(E, \mathcal{F})$ is a μ-system, there exists a feasible set $\{a_1, \dots, a_i, b_{j_{i+1}}, \dots, b_{j_k}\}$ for each $i \in \{0, 1, \dots, k-1\}$ such that $j_t \leq (t-1)\cdot\mu + 1$ $(t = i+1, \dots, k)$. As X_q is an optimal solution for w^q, we have $w^q(\{a_1, \dots, a_k\}) \geq w^q(\{a_1, \dots, a_i, b_{j_{i+1}}, \dots, b_{j_k}\}) \geq w^q(\{a_1, \dots, a_i, b_{i\cdot\mu+1}, \dots, b_{(k-1)\cdot\mu+1}\})$ and thus

$$w(a_{i+1})^q + \dots + w(a_k)^q \geq w(b_{i\cdot\mu+1})^q + \dots + w(b_{(k-1)\cdot\mu+1}) \quad (i = 0, \dots, k-1).$$

Hence, we have $w(a_k)^p + w(a_{k-1})^p + \dots + w(a_1)^p \geq w(b_{(k-1)\mu+1})^p + w(b_{(k-2)\mu+1})^p + \dots + w(b_1)^p$ by Karamata's inequality [12]. (Karamata's inequality is also known as Hardy–Littlewood–Pólya inequality [8].) Therefore, we obtain

$$w^p(X_q) = \sqrt[p]{\sum_{i=1}^k w(a_i)^p} \geq \sqrt[p]{\sum_{i=1}^k w(b_{(i-1)\mu+1})^p} \geq \sqrt[p]{\frac{1}{\mu}\sum_{i=1}^{\mu\cdot k} w(b_i)^p} = \frac{1}{\mu^{1/p}}\cdot w^p(X_p),$$

which proves the lemma. ◀

Finally, we show the upper bound for the case $p > q$ and $\mu^{1/q} \leq \alpha$.

▶ **Lemma 16.** *For any $p > q > 0$, $\mu^{1/q} \leq \alpha$, μ-system $(E, \mathcal{F})$ $(\mu \geq 1)$ and any weight $w : E \to \mathbb{R}_+$, we have $\eta_\alpha(E, \mathcal{F}, w; p, q) \leq 1$.*

Proof. By Lemma 15, there exists $X^* \in \arg\max\{w^p(X) : X \in \mathcal{F}\}$ such that $w^q(X^*) \geq \max\{w^q(X) : X \in \mathcal{F}\}/\mu^{1/q}$. Thus, we have $\eta_\alpha(E, \mathcal{F}, w; p, q) = 1$. ◀

Therefore, we get Theorem 13 by Lemmas 14, 15, and 16.

References

1 E. Anshelevich, A. Dasgupta, J. M. Kleinberg, É. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. *SIAM Journal on Computing*, 38(4):1602–1623, 2008.

2 A. Asadpour and A. Saberi. On the inefficiency ratio of stable equilibria in congestion games. In *Proceedings of WINE 2009*, pages 545–552. Springer, 2009.

3 Y. Azar, L. Epstein, Y. Richter, and G. J. Woeginger. All-norm approximation algorithms. *Journal of Algorithms*, 52(2):120–133, 2004.

4 F. J. Fabozzi, S. M. Focardi, and P. N. Kolm. *Quantitative Equity Investing: Techniques and Strategies*. Wiley, 2010.

5 R. Fujita, Y. Kobayashi, and K. Makino. Robust matchings and matroid intersections. *SIAM Journal on Discrete Mathematics*, 27:1234–1256, 2013.

6 A. Goel and A. Meyerson. Simultaneous optimization via approximate majorization for concave profits or convex costs. 44:301–323, 2006.

7 D. Golovin, A. Gupta, A. Kumar, and K. Tangwongsan. All-norms and all-l_p-norms approximation algorithms. In *Proceedings of IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science*, pages 199–210, 2008.

8 G. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, 1952.

9 R. Hassin and S. Rubinstein. Robust matchings. *SIAM Journal on Discrete Mathematics*, 15(4):530–537, 2002.

10 T. A. Jenkyns. The efficacy of the "greedy" algorithm. In *Proceedings of the 7th Southeastern Conference on Combinatorics, Graph Theory, and Computing, Utilitas Mathematica, Winnipeg*, pages 341–350, 1976.

11 N. Kakimura and K. Makino. Robust independence systems. *SIAM Journal on Discrete Mathematics*, 27(3):1257–1273, 2013.

12 J. Karamata. Sur une inégalité relative aux fonctions convexes. *Publications de l'Institut mathematique*, 1(1):145–147, 1932.

13 Y. Kawase and K. Makino. Nash equilibria with minimum potential in undirected broadcast games. *Theoretical Computer Science*, 482:33–47, 2013.

14 B. Korte and D. Hausmann. An analysis of the greedy heuristic for independence systems. *Algorithmic aspects of combinatorics*, 2:65–74, 1978.

15 B. Korte and J. Vygen. *Combinatorial Optimization: Theory and Algorithms*, volume 21. Springer, fifth edition.

16 A. Kumar and J. Kleinberg. Fairness measures for resource allocation. 36:657–680, 2006.

17 J. Mestre. Greedy in approximation algorithms. In *Proceedings of the 14th European Symposium on Algorithms*, pages 528–539, 2006.

18 D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

19 B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM Journal on Computing*, 24(2):227–234, 1995.

20 I. Rish and G. Grabarnik. *Sparse Modeling: Theory, Algorithms, and Applications.* CRC press, 2014.

21 B. Vöcking. Selfish load balancing. In N. Nisan, T. Roughgarden, É. Tardos, and V. V. Vazirani, editors, *Algorithmic Game Theory*, chapter 20, pages 517–542. Cambridge, 2007.