

February 1999
Revised June 1999

APPROXIMATE MAX-MIN RESOURCE SHARING
FOR STRUCTURED CONCAVE OPTIMIZATION*

M.D. Grigoriadis, L.G. Khachiyan, L. Porkolab and J. Villavicencio

LCSR-TR-374

Department of Computer Science
Hill Center for the Mathematical Sciences
Rutgers University
New Brunswick, NJ 08903.

*Research supported by the National Science Foundation under grant CCR-9618796.

APPROXIMATE MAX-MIN RESOURCE SHARING FOR STRUCTURED CONCAVE OPTIMIZATION*

M.D. GRIGORIADIS¹, L.G. KHACHIYAN¹, L. PORKOLAB², J. VILLAVICENCIO³

Abstract. We present a Lagrangian decomposition algorithm which uses logarithmic potential reduction to compute an ε -approximate solution of the general max-min resource sharing problem with M nonnegative concave constraints on a convex set B . We show that this algorithm runs in $O(M(\varepsilon^{-2} + \ln M))$ iterations, a data independent bound which is optimal up to polylogarithmic factors for any fixed relative accuracy $\varepsilon \in (0, 1)$. In the general structured case, B is the product of K convex blocks and each constraint function is block separable. For such models, an iteration of our method requires a $\Theta(\varepsilon)$ -approximate solution of K independent block maximization problems which can be computed in parallel.

AMS subject classification. 68Q25, 90C05, 90C27, 90C30, 90C06.

Key words. Approximation algorithm; Covering problem; Lagrangian decomposition; Logarithmic potential; Packing problem; Resource sharing; Structured optimization

1. Introduction. We consider the approximate solution of concave *max-min resource-sharing* problems of the form

$$(P) \quad \lambda^* = \max\{ \lambda \mid f(x) \geq \lambda e, x \in B \},$$

where $f : B \rightarrow \mathbb{R}_+^M$ is a given vector of M nonnegative continuous concave functions defined on a nonempty convex compact set B , e is the vector of all ones and, with no loss of generality, $\lambda^* > 0$. The problem is said to be *block-angular* whenever $B = B^1 \times \dots \times B^K$ for $K > 1$ given nonempty convex compact sets B^k , $k = 1, \dots, K$, and $f_m(x) = \sum_{k=1}^K f_m^k(x^k)$, where the $f_m^k(x^k)$, $m = 1, \dots, M$ are given continuous nonnegative concave functions for $x^k \in B^k$. We shall denote $\lambda(f) \doteq \min_{1 \leq m \leq M} f_m$ for any given $f \in \mathbb{R}_+^M$.

We shall be interested in computing an ε -*approximate solution* of this problem, i.e., for a given *relative tolerance* $\varepsilon \in (0, 1)$,

$$(P_\varepsilon) \quad \text{compute } x \in B \text{ that satisfies } f(x) \geq (1 - \varepsilon)\lambda^*e.$$

Our approach is based on the well-known duality relation:

$$(1.1) \quad \lambda^* = \max_{x \in B} \min_{p \in P} p^T f(x) = \min_{p \in P} \max_{x \in B} p^T f(x),$$

where $P \doteq \{p \in \mathbb{R}_+^M \mid e^T p = 1\}$. It follows that

$$(1.2) \quad \lambda^* = \min\{\Lambda(p) \mid p \in P\}, \quad (\text{Lagrangian dual})$$

where

$$(1.3) \quad \Lambda(p) = \max\{p^T f(x) \mid x \in B\}. \quad (\text{block problem})$$

*Research supported by the National Science Foundation under grant CCR-9618796.

¹Department of Computer Science, Busch Campus, Rutgers University, New Brunswick, NJ 08903.

²Max Planck Institut für Informatik, Saarbrücken, Germany.

³Mathematics Department, Catholic University of Chile, Santiago, Chile.

The exact optimality conditions for \mathcal{P} can thus be stated as follows: *A pair $x \in B$, $p \in P$ is optimal if and only if $\Lambda(p) = \lambda(f(x))$.*

In its simplest form, *Lagrangian* or *price-directive decomposition* is an iterative strategy that solves \mathcal{P} via its Lagrangian dual by computing a sequence of pairs p, x as follows. A *coordinator* uses the current $x \in B$ to compute some weights $p = p(f(x)) \in P$ corresponding to the coupling constraints $f(x) \geq \lambda e$, calls a *block solver* to compute a solution $\hat{x} \in B$ of (1.3) for this $p \in P$, and then makes a move from x to $(1 - \tau)x + \tau\hat{x}$ with an appropriate step length $\tau \in (0, 1]$. We call each such iteration a *coordination step*.

We shall only require an *approximate block solver* (\mathcal{ABS}), one that solves (1.3) to a given optimization tolerance $t > 0$, defined below.

$\mathcal{ABS}(p, t)$: compute $\hat{x} = \hat{x}(p) \in B$ such that $p^T f(\hat{x}) \geq (1 - t)\Lambda(p)$.

We shall eventually set $t = \Theta(\varepsilon)$ in our algorithm. For the block angular case, $\Lambda(p) = \sum_{k=1}^K \Lambda^k(p)$, where $\Lambda^k(p) \doteq \max\{p^T f^k(x^k) \mid x^k \in B^k\}$ and $\mathcal{ABS}(p, t)$ decomposes into K independent block solvers $\mathcal{ABS}^k(p, t)$, $k = 1, \dots, K$, each operating to the same accuracy $t > 0$. To simplify notation, we shall henceforth assume $K = 1$ and dispense with the superscript k .

By analogy to \mathcal{P}_ε , and based on the fact that λ^* is the optimal value of the Lagrangian dual (1.2), we define the ε -*approximate dual problem* as follows:

(\mathcal{D}_ε) compute $p \in P$ that satisfies $\Lambda(p) \leq (1 + \varepsilon)\lambda^*$.

We shall show below that, for a given relative accuracy $\varepsilon \in (0, 1)$, our proposed algorithm solves problems \mathcal{P}_ε and \mathcal{D}_ε in

$$(1.4) \quad N = O(M(\varepsilon^{-2} + \ln M))$$

coordination steps, each of which requires a call to $\mathcal{ABS}(p, \Theta(\varepsilon))$. On the other hand, it is easy to see that for $f(x) = x$ and $B = P$, any algorithm that solves \mathcal{P}_ε for $\varepsilon < 1$ by a sequence of block optimizations (1.3) must perform at least M such steps⁴. It follows that for a fixed ε , the bound (1.4) on the number of coordination steps is optimal to within a factor of $\ln(M)$.

The linear feasibility variant of \mathcal{P} , i.e., find $x \in B$ such that $f(x) = Ax \geq e$, often referred-to as the *fractional covering* problem, is solved in [5] by Lagrangian decomposition using exponential potential reduction. Up to polylogarithmic factors,

⁴This is true even for methods that use arbitrarily powerful coordinators which can compute p by taking into account previous iterates and block solutions as well as other data. See [2] for further discussion.

the number of iterations of that algorithm is proportional to M and ρ/ε^2 , where the data-dependent quantity $\rho = \max_m \max_{x \in B} f_m(x)$ is the *width* of B relative to $Ax \geq e$. The problem is further decomposed in [5] in a way that reduces this linear dependence on ρ down to $\log \rho$. However, this introduces additional constraints on the block problems which, in general, become NP-hard.

The logarithmic potential reduction algorithm for max-min optimization we present in this paper circumvents the issue of width altogether and uses the approximate block solvers \mathcal{ABS} on the original blocks (cf. [1] and [2]). We mention, in passing, that the first width-independent Lagrangian decomposition iteration bound for general min-max sharing was also based on a logarithmic potential function [2]. This bound was recently matched in [3] using an exponential potential reduction technique. It would be worthwhile to develop a width-independent exponential function reduction method for the max-min case as well.

The paper is organized as follows. In Section 2 we define the standard logarithmic barrier function and examine some of its properties. In Section 3 we develop our algorithm for solving \mathcal{P}_ε and \mathcal{D}_ε , and prove its correctness. Finally, we analyze the coordination complexity of the algorithm in Section 4. We will use the following notational abbreviations: $f \doteq f(x)$, $f' \doteq f(x')$, $\hat{f} \doteq f(\hat{x})$, $p \doteq p(f) \doteq p(f(x))$, for points $x, x', \hat{x} \in B$, respectively. The symbol e denotes the vector of all ones while e_i represents the i th unit vector.

2. Logarithmic potential function. We shall associate with the coupling inequalities $f \geq \lambda e$ the standard logarithmic potential function (see, e.g., Chapter 4 of [4]), of the form:

$$(2.1) \quad \Phi_t(\theta, f) = \ln \theta + \frac{t}{M} \sum_{m=1}^M \ln(f_m - \theta),$$

where $\theta \in \mathbb{R}$, $f = (f_1, f_2, \dots, f_M)$ are variables and t is a fixed positive parameter, identical to that used for $\mathcal{ABS}(p, t)$. The function Φ_t is well-defined for $0 < \theta < \lambda(f) = \min\{f_1, f_2, \dots, f_M\}$. This implies that $f \in \text{Re}_{++}^M$, which will be the case for all iterates of the algorithm we shall present in Section 3.

Similar to [2, 6], we define the *reduced potential function* as the maximum of $\Phi_t(\theta, f)$ over $\theta \in (0, \lambda(f))$ for a fixed $f \in \mathbb{R}_{++}^M$, i.e.,

$$(2.2) \quad \phi_t(f) = \max_{0 < \theta < \lambda(f)} \Phi_t(\theta, f).$$

The maximizer $\theta(f)$ of $\Phi_t(\theta, f)$ can be determined from the first-order optimality condition:

$$(2.3) \quad \frac{t\theta}{M} \sum_{m=1}^M \frac{1}{f_m - \theta} = 1,$$

which has a unique root since its left side is a strictly increasing function of θ . We can thus write the reduced potential function as:

$$\phi_t(f) = \Phi_t(\theta(f), f).$$

It is easy to see that the smooth function $\theta(f)$ approximates the piecewise nonlinear concave function $\lambda(f)$ as follows:

$$\frac{\lambda(f)}{1+t} \leq \theta(f) \leq \frac{\lambda(f)}{1+t/M},$$

a property which motivates our approach.

Next, we define the *logarithmic dual vector* $p \doteq p(f)$ for a fixed $f \in \mathbb{R}_{++}^M$ to be:

$$(2.4) \quad p_m(f) = \frac{t}{M} \frac{\theta(f)}{f_m - \theta(f)}, \quad m = 1, \dots, M,$$

where $p(f) \in P$ by (2.3). A useful consequence of this definition is the following identity.

PROPOSITION 1. $p(f)^T f = (1+t)\theta(f)$.

Proof. Denoting $\theta \doteq \theta(f)$, we write

$$p^T f = \frac{t\theta}{M} \sum_{m=1}^M \frac{f_m}{f_m - \theta} = \frac{t\theta}{M} \sum_{m=1}^M \left(1 + \frac{\theta}{f_m - \theta} \right) = t\theta + \theta \sum_{m=1}^M p_m = (1+t)\theta. \quad \square$$

A more important observation is that the accuracy with which the optimality criteria $\Lambda(p) = \lambda(f)$ are to be met at a given point $x \in B$ can be approximated by the quantity:

$$(2.5) \quad \nu \doteq \nu(x, \hat{x}) = \frac{p^T \hat{f} - p^T f}{p^T \hat{f} + p^T f}$$

where $p \in P$ from (2.4), $f = f(x)$ and $\hat{f} = f(\hat{x})$ for an approximate block solution $\hat{x} \in B$ produced by $\mathcal{ABS}(p, t)$. The theorem below states that a pair x, p solves \mathcal{P}_ε and \mathcal{D}_ε , respectively, whenever ν and t are of order ε .

LEMMA 1. *Suppose $\varepsilon \in (0, 1)$ and $t = \varepsilon/6$. For a given point $x \in B$, let $p \in P$ be computed by (2.4) and \hat{x} computed by $\mathcal{ABS}(p, t)$. If $\nu(x, \hat{x}) \leq t$, then the pair x, p solves \mathcal{P}_ε and \mathcal{D}_ε , respectively.*

Proof. Use (2.5) to rewrite the condition $\nu \leq t$ as follows:

$$(1-t)p^T \hat{f} \leq (1+t)p^T f.$$

Since $p^T \hat{f} \geq \Lambda(p)(1-t)$, $p^T f = (1+t)\theta$ by Proposition 1 and $\theta = \theta(f) < \lambda(f)$, we obtain:

$$(2.6) \quad \Lambda(p) \leq \frac{1+t}{(1-t)^2} p^T f \leq \left(\frac{1+t}{1-t} \right)^2 \lambda(f) \leq (1+\varepsilon)\lambda(f),$$

where the last inequality follows from the assumption $t = \varepsilon/6$. Given that $\Lambda(p) \geq \lambda^*$, we have $\lambda^* \leq (1+\varepsilon)\lambda(f) \leq \lambda(f)/(1-\varepsilon)$, which is \mathcal{P}_ε . On the other hand, $\lambda(f) \leq \lambda^*$, so that (2.6) implies $\Lambda(p) \leq (1+\varepsilon)\lambda^*$, which is \mathcal{D}_ε . \square

REMARK 1. The bound of Lemma 1 is close to the best possible: for $\varepsilon \in (0, 1/2]$ we have $\nu \leq \varepsilon/4$ for any pair $x \in B$, $p \in P$ that solves \mathcal{P}_ε and \mathcal{D}_ε , respectively. To see this, first consider that

$$\Lambda(p) \leq (1+\varepsilon)\lambda^* \leq \frac{1+\varepsilon}{1-\varepsilon} \lambda(f),$$

which simplifies to $\Lambda(p) \leq (1+4\varepsilon)\lambda(f)$ for $\varepsilon \in (0, 1/2]$. Then, use this inequality, (2.5), and the fact that $\lambda(f) \leq p^T f$, to write:

$$\nu \leq \frac{\Lambda(p) - p^T f}{p^T \hat{f} + p^T f} \leq \frac{4\varepsilon p^T f}{p^T \hat{f} + p^T f} \leq 4\varepsilon. \quad \square$$

3. The approximation algorithm. We shall now state our Algorithm \mathcal{A} to compute solutions of both problems \mathcal{P}_ε and \mathcal{D}_ε , as a direct implementation of the Lagrangian decomposition scheme stated in the Introduction. The algorithm accepts as input f, B, ε and an initial point $x = x^0 \doteq \frac{1}{M} \sum_{m=1}^M \hat{x}^{(m)} \in B$, computed by $x^{(m)} := \text{ABS}(e_m, 1/2)$.

Algorithm $\mathcal{A}(f, B, \varepsilon, x)$:

$t := \varepsilon/6$

while $\nu > t$ **do**

 Compute $\theta(f)$ from (2.3) and $p \in P$ from (2.4).

$\hat{x} := \text{ABS}(p, t)$.

 Compute $\nu \doteq \nu(x, \hat{x})$ from (2.5).

$x := (1-\tau)x + \tau\hat{x}$, for an appropriate step length $\tau \in (0, 1]$.

end

return(x, p)

Our subsequent analysis uses the step length

$$(3.1) \quad \tau = \frac{t\theta\nu}{2M(p^T \hat{f} + p^T f)},$$

which is strictly feasible, i.e., $\tau \in (0, 1)$. (To see this, substitute $t = \varepsilon/6$ in (3.1) and use the inequality $\theta/(p^T \hat{f} + p^T f) \leq 1$, which is a straightforward consequence of

Proposition 1.) In practice, one usually computes τ by performing a line search to maximize $\phi_t(x + \tau(\hat{x} - x))$. Our analysis remains valid for such step lengths.

Algorithm \mathcal{A} is correct since $\tau \in (0, 1)$ as indicated above and since by Lemma 1, the pair $x \in B$, $p \in P$ solves \mathcal{P}_ε and \mathcal{D}_ε , respectively, when the algorithm halts.

4. Analysis of the approximation algorithm. Our next task is to derive the iteration bound (1.4) for our algorithm as claimed in the Introduction. We shall first establish several observations. In Lemma 2 we bound the error in the initial approximation $x^0 \in B$ as defined in Section 3. Lemma 3 shows that each coordination step achieves a sizable guaranteed increase in the value of the reduced potential function $\phi_t(f)$. In contrast, Lemma 4 bounds the sum of such increases between any two, not necessarily consecutive, iterates. These results ultimately lead us to the iteration bound stated in Theorem 1, which is further improved by a factor $O(\varepsilon^{-1})$ by employing a simple error-scaling technique akin to that used in [6] for structured min-max problems.

LEMMA 2. $\lambda^* \leq \Lambda(p) \leq 2Mp^T f(x^0)$ for all $p \in P$.

Proof. The left inequality is from (1.2)-(1.3). To show the right inequality, note that for any $p \in P$, (1.3) provides:

$$\Lambda(p) = \max\{p^T f(x) \mid x \in B\} \leq \sum_{m=1}^M p_m \max\{f_m(x) \mid x \in B\} = \sum_{m=1}^M p_m \Lambda(e_m).$$

Now, $\Lambda(e_m) \leq 2f_m(\hat{x}^{(m)})$, where $\hat{x}^{(m)}$ is the approximate block solution computed by $\mathcal{ABS}(e_m, 1/2)$. Then, by using the concavity of the nonnegative functions f_m , we obtain:

$$f_m(\hat{x}^{(m)}) \leq \sum_{\ell=1}^M f_m(\hat{x}^{(\ell)}) \leq M f_m \left(\frac{1}{M} \sum_{\ell=1}^M \hat{x}^{(\ell)} \right) = M f_m(x^0). \quad \square$$

Next, we prove that the increase in the reduced potential $\phi_t(f)$ is sufficiently large at each iteration.

LEMMA 3. For any two consecutive iterates x, x' of Algorithm \mathcal{A} :

$$\phi_t(f') - \phi_t(f) \geq t\nu^2/4M.$$

Proof. From Algorithm \mathcal{A} we have $x' = (1 - \tau)x + \tau\hat{x}$. Denote $\Lambda \doteq \Lambda(p(f))$. By the concavity of the f_m and definition (2.4),

$$\begin{aligned} (4.1) \quad f'_m - \theta &\geq (1 - \tau)f_m + \tau\hat{f}_m - \theta = (f_m - \theta) \left(1 + \tau \frac{\hat{f}_m - f_m}{f_m - \theta} \right) \\ &= (f_m - \theta) \left(1 + \frac{\tau M}{t\theta} p_m(\hat{f}_m - f_m) \right). \end{aligned}$$

In order to bound the last expression above, consider that by definition (3.1),

$$\left| \frac{\tau M}{t\theta} p_m(\hat{f}_m - f_m) \right| \leq \frac{\tau M}{t\theta} p_m(\hat{f}_m + f_m) \leq \frac{\tau M}{t\theta} (p^T \hat{f} + p^T f) = \frac{\nu}{2} \leq \frac{1}{2}.$$

Accordingly, (4.1) gives $f'_m - \theta > 0$, $m = 1, \dots, M$, so that $\lambda(f') > \theta$. From the definition (2.2) of $\phi_t(f')$,

$$(4.2) \quad \phi_t(f') = \max_{0 < \xi < \lambda(f')} \Phi_t(\xi, f') \geq \Phi_t(\theta, f') = \ln \theta + \frac{t}{M} \sum_{m=1}^M \ln(f'_m - \theta),$$

where θ denotes the root $\theta(f)$ of (2.3). By combining inequalities (4.1) and (4.2), and by using the definition of $\phi_t(f) \equiv \Phi(\theta(f), f)$, we obtain:

$$(4.3) \quad \phi_t(f') \geq \phi_t(f) + \frac{t}{M} \sum_{m=1}^M \ln \left(1 + \frac{\tau M}{t\theta} p_m(\hat{f}_m - f_m) \right).$$

We now use the inequality $\ln(1+z) \geq z - z^2$, for all $z \geq -1/2$, to write:

$$\phi_t(f') - \phi_t(f) \geq \tau \frac{p^T \hat{f} - p^T f}{\theta} - \frac{\tau^2 M}{\theta^2 t} \|D(f - \hat{f})\|^2,$$

where $D = \text{diag}(p_1, \dots, p_m)$ and $\|\cdot\|$ denotes the 2-norm. Furthermore, $\|D(f - \hat{f})\| \leq \|D(f + \hat{f})\|_1 = p^T f + p^T \hat{f}$, which implies that

$$\phi_t(f') - \phi_t(f) \geq \tau \frac{p^T \hat{f} - p^T f}{\theta} - \frac{\tau^2 M}{\theta^2 t} (p^T f + p^T \hat{f})^2.$$

This last inequality proves the claim for the value of τ given by (3.1). \square

In contrast to the previous lemma, our third observation provides that the increase in the reduced potential $\phi_t(f)$ cannot be too large even after an arbitrary number of iterations.

LEMMA 4. *For any $x \in B$ such that $\lambda(f) > 0$ and another point $x' \in B$:*

$$\phi_t(f') - \phi_t(f) \leq (1+t) \ln \frac{\Lambda(p)}{p^T f},$$

where p is defined by (2.4).

Proof. Denote $\theta \doteq \theta(f)$, $\theta' \doteq \theta(f')$ and $\Lambda \doteq \Lambda(p) = \max\{p^T f(x) \mid x \in B\}$ as defined in (1.3). Assume with no loss of generality that $\lambda(f') > 0$, since $\phi_t(f') = -\infty$ for any $x' \in B$ for which $\lambda(f') = 0$. We can thus write:

$$\begin{aligned} \phi_t(f') - \phi_t(f) &= \ln \frac{\theta'}{\theta} + \frac{t}{M} \sum_{m=1}^M \ln \left(\frac{f'_m - \theta'}{f_m - \theta} \right) \\ &= \ln \frac{\theta'}{\theta} + \frac{t}{M} \sum_{m=1}^M \ln \left(\frac{M}{t\theta} p_m (f'_m - \theta') \right) \\ &= \ln \frac{\theta'}{\theta} + t \ln \frac{M}{t\theta} + \frac{t}{M} \sum_{m=1}^M \ln (p_m (f'_m - \theta')). \end{aligned}$$

Next, using the concavity of $\ln(\cdot)$ in the last expression we obtain:

$$\begin{aligned} \phi_t(f') - \phi_t(f) &\leq \ln \frac{\theta'}{\theta} + t \ln \frac{M}{t\theta} + t \ln \left(\frac{1}{M} p^T (f' - \theta' e) \right) \\ &\leq \ln \frac{\theta'}{\theta} + t \ln \frac{1}{t\theta} + t \ln (\Lambda - \theta'), \end{aligned}$$

which is further simplified as follows:

$$\begin{aligned} \phi_t(f') - \phi_t(f) &\leq \max_{\xi \in (0, \Lambda)} \left\{ \ln \frac{\xi}{\theta} + t \ln \frac{1}{t\theta} + t \ln (\Lambda - \xi) \right\} \\ &= (1+t) \ln (\Lambda / (1+t)\theta) = (1+t) \ln (\Lambda / p^T f) \quad \square \end{aligned}$$

We are now in a position to address the coordination complexity of Algorithm \mathcal{A} by combining the lower and upper bounds we have thus far obtained for the increase in $\phi_t(f)$ from one iteration to the next.

THEOREM 1. *For any given relative accuracy $\varepsilon > 0$, Algorithm \mathcal{A} solves problems \mathcal{P}_ε and \mathcal{D}_ε in*

$$N = O(M(\varepsilon^{-1} \ln M + \varepsilon^{-2}))$$

coordination steps.

Proof. First, let N_0 be the number of iterations of Algorithm \mathcal{A} required to obtain an iterate x^1 with a corresponding optimality error $\nu \leq 1/2$, starting from the initial point x^0 . For as long as $\nu > 1/2$, each iteration increases the reduced potential by at least $t/16M$ (Lemma 3). However, by Lemma 4, the total increase in the value of the reduced potential can be bounded as follows:

$$(4.4) \quad \phi_t(f^1) - \phi_t(f^0) \leq (1+t) \ln (\Lambda(p^0) / p^{0T} f^0).$$

Since $t = \varepsilon/6$ and $\Lambda(p^0) \leq 2Mp^{0T}f^0$ by Lemma 2, it follows that $N_0 = O(\varepsilon^{-1}M \ln M)$.

Next, suppose that the error is $\nu_\ell \leq 1/2^\ell$ for some iterate $x^\ell \in B$ and let N_ℓ be the number of iterations required to halve this error, for $\ell = 1, 2, \dots$. Again, Lemma 3 provides:

$$(4.5) \quad \phi_t(f^{\ell+1}) - \phi_t(f^\ell) \geq N_\ell t\nu_\ell^2/16M.$$

To bound the left side of this inequality, consider:

$$(1 - \nu_\ell)p^{\ell T}\hat{f}^\ell = (1 + \nu_\ell)p^{\ell T}f^\ell,$$

directly from the definition of ν_ℓ in (2.5). And since $p^{\ell T}\hat{f}^\ell \geq (1-t)\Lambda(p^\ell)$ for $\mathcal{ABS}(p^\ell, t)$, we have:

$$\frac{\Lambda(p^\ell)}{p^{\ell T}f^\ell} \leq \frac{1 + \nu_\ell}{(1-t)(1-\nu_\ell)}.$$

This inequality, along with the fact that $t \leq \nu_\ell \leq 1/2$, implies:

$$\frac{\Lambda(p^\ell)}{p^{\ell T}f^\ell} \leq \frac{1 + \nu_\ell}{(1-\nu_\ell)^2} \leq 1 + 10\nu_\ell.$$

Now, Lemma 4 for $x' \doteq x^{\ell+1}$ and $x \doteq x^\ell$ provides:

$$\phi_t(f^{\ell+1}) - \phi_t(f^\ell) \leq (1+t)\ln(1+10\nu_\ell) \leq 10(1+t)\nu_\ell,$$

which, together with (4.5), results in the bound $N_\ell = O(M/t\nu_\ell)$. The total number of coordination steps N in the claim is obtained by summing the N_ℓ over $\ell = 0, 1, \dots, \lceil \lg(1/\varepsilon) \rceil$. \square

The coordination complexity of Algorithm \mathcal{A} given by Theorem 1 is for a fixed value of the parameter t . The algorithm can be implemented and its coordination complexity improved by embedding Algorithm \mathcal{A} within a sequence of *scaling phases* that gradually reduce t to the desired accuracy, much like implementations of path-following methods for convex programs. The s th scaling phase sets $\varepsilon_s := \varepsilon_{s-1}/2$, correspondingly $t_s := \varepsilon_s/6$, and uses the current approximate point x^{s-1} as its initial solution. The entire scaling algorithm is initialized (for phase $s = 0$) with the same initial point $x_0 \in B$ as before, which gives $\varepsilon_0 = 1 - 1/2M$. The resulting coordination complexity of the overall scheme is analyzed below.

THEOREM 2. *For any given relative accuracy $\varepsilon > 0$, the error scaling implementation of Algorithm \mathcal{A} computes solutions x, p of problems \mathcal{P}_ε and \mathcal{D}_ε , respectively, in*

$$N = O(M(\ln M + \varepsilon^{-2}))$$

coordination steps.

Proof. Denote by \mathcal{N}_s the number of coordination steps in the $s+1$ st scaling phase, $s = 0, 1, \dots$. By Theorem 1, $\mathcal{N}_0 = O(M \ln M)$. It remains to show that $\mathcal{N}_s = O(M/\varepsilon_s^2)$ for each subsequent scaling phase.

Our arguments are analogous to those used in the proof of Theorem 1. By Lemma 3, each iteration of the s th scaling phase increases the (current) reduced potential function by $t_s \nu_s^2 / 4M \geq t_s^3 / 4M = \Omega(\varepsilon_s^3 / M)$.

By invoking Lemma 4 with $x \doteq x^s$ and $x' \doteq x^{s+1}$, the total increase of the potential in the s th phase can be bounded by:

$$(4.6) \quad \phi_{t_s}(f^{s+1}) - \phi_{t_s}(f^s) \leq (1 + \varepsilon_s/6) \ln(\Lambda^s / p^{sT} f^s).$$

Furthermore, since x^s is a $2\varepsilon_s$ -approximate solution of problem \mathcal{P} ,

$$(4.7) \quad \Lambda^s \leq \frac{1 + 2\varepsilon_s}{1 - 2\varepsilon_s} \lambda(f^s) \leq (1 + 8\varepsilon_s) p^{sT} f^s.$$

The bound $\mathcal{N}_s = O(M/\varepsilon_s^2)$ is deduced from (4.7) and the fact that $\ln(1 + 8\alpha) \leq 8\alpha$ for all $\alpha > 0$.

As before, the overall coordination complexity is obtained by adding the coordination bounds \mathcal{N}_s for all scaling phases. \square

5. References

- [1] M.D. GRIGORIADIS AND L.G. KHACHIYAN, *Fast approximation schemes for convex programs with many blocks and coupling constraints*, SIAM Journal on Optimization (1994), 86-103.
- [2] ———, *Coordination complexity of parallel price-directive decomposition*, Mathematics of Operations Research (1996), 321-340.
- [3] N. GARG AND J. KÖNEMANN, *Faster and simpler algorithms for multicommodity flow and other fractional packing problems*, Proceedings of the 39th Annual Symp. on Foundations of Comp. Sci. 300-309, Palo Alto, CA, November 1998. IEEE.
- [4] Y. NESTEROV AND A. NEMIROVSKII, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.
- [5] S.A. PLOTKIN, D.B. SHMOYS AND E. TARDOS, *Fast approximation algorithms for fractional packing and covering problems*, Mathematics of Operations Research (1995), 257-301.
- [6] J. VILLAVICENCIO AND M.D. GRIGORIADIS, *Approximate Lagrangian decomposition with a modified Karmarkar logarithmic potential*, in Network Optimization, P. Pardalos, D.W. Hearn and W.W. Hager, Eds, Lecture Notes in Economics and Mathematical Systems 450, Springer-Verlag, Berlin, 1997, pp. 471-485.