CENTER FOR
PARALLEL OPTIMIZATION

PARALLEL CONSTRAINT DISTRIBUTION IN CONVEX QUADRATIC PROGRAMMING

by

Michael C. Ferris

Computer Sciences Technical Report #1009

February 1991

Parallel Constraint Distribution in Convex Quadratic Programming*

Michael C. Ferris†

February 1991

Abstract. We consider convex quadratic programs with large numbers of constraints. We distribute these constraints among several parallel processors and modify the objective function for each of these subproblems with Lagrange multiplier information from the other processors. New Lagrange multiplier information is aggregated in a master processor and the whole process is repeated. Linear convergence is established for strongly convex quadratic programs by formulating the algorithm in an appropriate dual space. The algorithm corresponds to a step of an iterative matrix splitting algorithm for a symmetric linear complementarity problem followed by a projection onto a subspace.

Key words. Parallel Optimization, Augmented Lagrangians, Quadratic Programs, Matrix Splitting, Linear Convergence

Abbreviated title. Parallel Constraint Distribution

1 Introduction

We are concerned with the problem

$$\begin{array}{ll} \text{minimize} & c^T x + \frac{1}{2} x^T Q x \\ \text{subject to} & A_1 x \leq a_1, \ldots, A_p x \leq a_p \end{array}$$

where $c \in \mathbb{R}^n$, $Q \in \mathbb{R}^{n \times n}$, $A_i \in \mathbb{R}^{m_i \times n}$, $a_i \in \mathbb{R}^{m_i}$ and Q is symmetric and positive definite. Our principal aim is to distribute the p constraint blocks among p parallel processors together with an appropriately modified objective function. We then solve each of these p subproblems independently, aggregate Lagrange multiplier information from the processors and repeat. The method we describe here is closely related to the one given in [2]. References to other

*This material is based on research supported by the Air Force Office of Scientific Research Grant AFOSR-89-0410

†Computer Sciences Department, University of Wisconsin, 1210 West Dayton Street, Madison, Wisconsin 53706

constraint distribution algorithms can be found in that paper. The key to our approach lies in the precise form of the modified objective function to be optimized by each processor. The modified objectives are made up of the original objective function plus augmented Lagrangian terms involving the constraints handled by the other processors.

In this paper, we show that under the assumption of a strongly convex quadratic objective and linear independence of each of the distributed constraint blocks, the parallel constraint distribution (PCD) algorithm converges linearly from any starting point for a feasible problem. The key to the convergence proof is to show that in the dual space, an iteration of the proposed parallel constraint distribution algorithm is equivalent to a step of an iterative matrix splitting method for a symmetric linear complementarity problem followed by a subspace projection.

A word about our notation now. For a vector x in the n-dimensional real space $\mathbb{R}^n$, x_+ will denote the vector in $\mathbb{R}^n$ with components $(x_+)_i := \max\{x_i, 0\}$, $i = 1, \ldots, n$. The standard inner product of $\mathbb{R}^n$ will be denoted either by $\langle x, y\rangle$ or $x^T y$. The Euclidean or 2–norm $(x^T x)^{\frac{1}{2}}$, will be denoted by $\|\cdot\|$. For an $m \times n$ real matrix A, signified by $A \in \mathbb{R}^{m\times n}$, A^T will denote the transpose. The identity matrix of any order will be given by I. The nonnegative orthant in $\mathbb{R}^n$ will be denoted by $\mathbb{R}^n_+$. The notation $P(x \mid C)$ will be used to define the projection of the point x onto the closed convex set C.

2 Parallel constraint distribution for quadratic programs

For simplicity we consider a quadratic program with 3 blocks of inequality constraints. Routine extension to p blocks can be achieved by appropriate extension and permutation of subscripts. Equality constraints can also be incorporated in an straightforward manner. Consider then the problem

$$\begin{array}{ll} \text{minimize} & c^T x + \frac{1}{2} x^T Q x \\ \text{subject to} & A_l x \le a_l,\ l = 1, 2, 3 \end{array} \tag{1}$$

where $c \in \mathbb{R}^n$, $Q \in \mathbb{R}^{n\times n}$, $A_l \in \mathbb{R}^{m_l \times n}$, $a_l \in \mathbb{R}^{m_l}$ and Q is symmetric and positive definite. We now describe the algorithm in detail. At iteration i we distribute the constraints of (1) among 3 parallel processors ($l = 1, 2, 3$) as follows

$$\begin{array}{ll} \underset{x_l}{\text{minimize}} & c^T x_l + \frac{1}{2} x_l^T Q x_l + \frac{1}{2\gamma}\left[\sum_{\substack{k=1\\k\neq l}}^{3} \left\|(\gamma(A_k x_l - a_k) + t^i_{kl})_+\right\|^2\right] \\ \text{subject to} & A_l x_l \le a_l \end{array} \tag{2}$$

where γ is a positive number and t^i_{kl} are estimates of the Lagrange multipliers from the previous iteration. We note that the subproblems (2) of the algorithm split the constraints of the original quadratic program (1) between them in the form of split explicit constraints as well as augmented Lagrangian terms involving the other constraints. The principal objective that has been achieved is that the **explicit** constraints of each of the subproblems are a subset of the constraints of the original problem. The difference between this algorithm and standard augmented Lagrangian methods (see [1, 9, 10]) is that the multiplier update is carried out explicitly rather than with the traditional gradient updating scheme. Each subproblem is then solved and a point $(\bar{x}_l^{i+1}, \bar{s}_l^{i+1}) \in \mathbb{R}^{n+m_l}$, $l = 1, 2, 3$, which satisfies the Karush–Kuhn–Tucker conditions [5] for subproblems (2) is obtained. We define

$$\bar{t}_{jl}^{i+1} = \left(\gamma(A_j\bar{x}_l^{i+1} - a_j) + t_{jl}^i\right)_+, j \neq l \tag{3}$$

and let

$$s_l^{i+1} = \frac{\bar{s}_l^{i+1} + \sum_{\substack{k=1 \\ k\neq l}}^3 \bar{t}_{kl}^{i+1}}{3} \tag{4}$$

and

$$t_{jl}^{i+1} = s_j^{i+1}, j \neq l \tag{5}$$

This completes one iteration of the PCD algorithm. Clearly, steps (2), (3) and (5) should be executed in parallel, while step (4) should be executed on a master processor. For completeness, we give the algorithm below:

PCD Algorithm

Initialization: Start with any s_l^0, $l = 1, 2, 3$.

Parallel iteration: In parallel, $(l = 1, 2, 3)$, perform the following steps.
Having s_l^i compute:

1. t_{jl}^i using (5)
2. $(\bar{x}_l^{i+1}, \bar{s}_l^{i+1})$ using (2)
3. $\bar{t}_{jl}^{i+1})$ using (3)

Synchronization: evaluate s_l^{i+1}, $l = 1, 2, 3$ using (4). If converged, then stop, else return to parallel iteration.

In the sequel we will show that s_l^i, $l = 1, 2, 3$ converge linearly to a set of optimal multipliers for the dual of (1). Furthermore, we show that the solutions of the subproblems $\bar{x}_l^i$ each converge to the optimal solution of (1). In order to facilitate this we will rewrite the iterates of the algorithm exclusively in the dual space. We shall use the following easily established result frequently throughout the paper, so we state it as a lemma.

Lemma 1 *Let $b, d \in \mathbb{R}^n$. Then*

$$b = d_+ \iff b - d \geq 0,\ b^T(b - d) = 0,\ b \geq 0$$

It follows from Lemma 1 that $(\bar{x}_l^{i+1}, \bar{s}_l^{i+1})$ satisfy the following Karush–Kuhn–Tucker conditions

$$\begin{array}{r} c + Q\bar{x}_l^{i+1} + \sum_{\substack{k=1\\k\neq l}}^3 A_k^T\left(\gamma(A_k\bar{x}_l^{i+1} - a_k) + t_{kl}^i\right)_+ + A_l^T\bar{s}_l^{i+1} = 0 \\ \bar{s}_l^{i+1} = \left(\bar{s}_l^{i+1} + \gamma(A_l\bar{x}_l^{i+1} - a_l)\right)_+ \end{array} \qquad l = 1,2,3 \tag{6}$$

or equivalently

$$\begin{aligned} \bar{x}_l^{i+1} &= -Q^{-1}(c + \textstyle\sum_{\substack{k=1\\k\neq l}}^3 A_k^T\bar{t}_{kl}^{i+1} + A_l^T\bar{s}_l^{i+1}) \\ \bar{s}_l^{i+1} &= \left(\bar{s}_l^{i+1} + \gamma(A_l\bar{x}_l^{i+1} - a_l)\right)_+ \qquad l = 1,2,3,\ j = 1,2,3,\ j \neq l \\ \bar{t}_{jl}^{i+1} &= \left(\gamma(A_j\bar{x}_l^{i+1} - a_j) + t_{jl}^i\right)_+ \end{aligned} \tag{7}$$

We use the first equation of (7) to eliminate $\bar{x}_l^{i+1}$ from the second and third equations of (7). This leads to the following system

$$\begin{aligned} \bar{s}_l^{i+1} &= \left(\bar{s}_l^{i+1} - (\gamma A_lQ^{-1}A_l^T\bar{s}_l^{i+1} + \gamma A_lQ^{-1}\textstyle\sum_{\substack{k=1\\k\neq l}}^3 A_k^T\bar{t}_{kl}^{i+1} + \gamma(A_lQ^{-1}c + a_l))\right)_+ \\ \bar{t}_{jl}^{i+1} &= \left(t_{jl}^{i+1} - (t_{jl}^i + \gamma A_jQ^{-1}A_l^T\bar{s}_l^{i+1} + \gamma A_jQ^{-1}\textstyle\sum_{\substack{k=1\\k\neq l}}^3 A_k^T\bar{t}_{kl}^{i+1} + \gamma(A_jQ^{-1}c + a_j))\right)_+ \end{aligned} \tag{8}$$

for $l = 1,2,3$ and $j = 1,2,3$ with $j \neq l$. Let us define new variables in blocks as follows

$$\begin{aligned} u_1^i &= (s_1^i, t_{21}^i, t_{31}^i) \\ u_2^i &= (t_{12}^i, s_2^i, t_{32}^i) \\ u_3^i &= (t_{13}^i, t_{23}^i, s_3^i) \end{aligned} \tag{9}$$

(with a similar notation for the barred variables). We can then rewrite (8) as follows

$$\bar{u}_l^{i+1} = \left(\bar{u}_l^{i+1} - ((H + J_l)\bar{u}_l^{i+1} - J_lu_l^i + h)\right)_+$$

for $l = 1,2,3$, where

$$H = \gamma\begin{bmatrix} A_1Q^{-1}A_1^T & A_1Q^{-1}A_2^T & A_1Q^{-1}A_3^T \\ A_2Q^{-1}A_1^T & A_2Q^{-1}A_2^T & A_2Q^{-1}A_3^T \\ A_3Q^{-1}A_1^T & A_3Q^{-1}A_2^T & A_3Q^{-1}A_3^T \end{bmatrix}, \qquad h = \gamma\begin{bmatrix} A_1Q^{-1}c + a_1 \\ A_2Q^{-1}c + a_2 \\ A_3Q^{-1}c + a_3 \end{bmatrix} \tag{10}$$

and J_l is defined by

$$J_1 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \qquad J_2 := \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I \end{bmatrix} \qquad J_3 := \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where the blocks are partitioned as in the definition of H. If we let

$$
\begin{aligned}
\bar{z}^{i+1} &= (\bar{u}_1^{i+1}, \bar{u}_2^{i+1}, \bar{u}_3^{i+1}) \\
z^i &= (u_1^i, u_2^i, u_3^i)
\end{aligned} \tag{11}
$$

and invoke Lemma 1, the following symmetric linear complementarity problem in the variable $\bar{z}^{i+1}$ ensues

$$
B\bar{z}^{i+1} + Cz^i + q \geq 0, \ \left\langle \bar{z}^{i+1}, B\bar{z}^{i+1} + Cz^i + q \right\rangle = 0, \ \bar{z}^{i+1} \geq 0 \tag{12}
$$

Here

$$
B = \begin{bmatrix} H + J_1 & 0 & 0 \\ 0 & H + J_2 & 0 \\ 0 & 0 & H + J_3 \end{bmatrix} \tag{13}
$$

$$
C = \begin{bmatrix} -J_1 & 0 & 0 \\ 0 & -J_2 & 0 \\ 0 & 0 & -J_3 \end{bmatrix} \tag{14}
$$

and q is given by

$$
q = \begin{bmatrix} h \\ h \\ h \end{bmatrix} \tag{15}
$$

We observe that due to the structure of B and C, (12) can be performed on 3 independent processors (which corresponds to the 3 subproblems in (2)). However, there is no coupling between the subproblems and this will produce poor convergence. We therefore modify the matrix splitting algorithm in order to produce a simple coupling between processors.

We remark that the matrices B and C constitute a splitting of a matrix M given by

$$
M = \begin{bmatrix} H & 0 & 0 \\ 0 & H & 0 \\ 0 & 0 & H \end{bmatrix} \tag{16}
$$

Note that if our original quadratic program (1) is feasible, then it it solvable. Hence its Wolfe dual is solvable and this dual is LCP(H,h), given by

$$
Hu + h \geq 0, \ \langle u, Hu + h \rangle = 0, \ u \geq 0
$$

The matrix splitting step given by (12) is a splitting for LCP(M,q)

$$
Mz + q \geq 0, \ \langle z, Mz + q \rangle = 0, \ z \geq 0 \tag{17}
$$

where M is defined in (16) and q in (15). In fact LCP(M,q) constitutes a replication of the Wolfe dual LCP(H,h), 3 times. Thus the solvability of LCP(H,h) is equivalent to the solvability of LCP(M,q). It is therefore clear that the solution set of LCP(M,q) is given by

$$Z^* = U^* \times U^* \times U^*$$

where U^* is the solution set of LCP(H,h). We now construct an algorithm which has two steps. The first step is an iteration of the matrix splitting algorithm

$$\bar{z}^{i+1} = \left(\bar{z}^{i+1} - (B\bar{z}^{i+1} + Cz^i + q)\right)_+ \tag{18}$$

as described in (12), where we require the splitting (B,C) to be regular, that is

$$M = B + C, \quad B - C \text{ is positive definite} \tag{19}$$

The second step is used to force the elements of z^i to converge to a point in Z^* where $u_1 = u_2 = u_3$ and thus introduce a very simple coupling between the subproblems. Hence we define a subspace L by

$$L := \{z = (u_1, u_2, u_3) \mid u_1 = u_2 = u_3\}$$

and generate z^{i+1} by projecting the iterate given by the matrix splitting step onto L, that is

$$z^{i+1} = P(\bar{z}^{i+1} \mid L) \tag{20}$$

This exactly corresponds to the master processor step described in (4).

We now proceed to analyze the algorithm. We will invoke the following merit function to prove linear convergence in the dual space

$$f(z) := \frac{1}{2} z^T M z + q^T z \tag{21}$$

Our main theorem will require the following result due to Tseng and Luo[11, Theorem 2.1] which we state here for completeness.

Proposition 2 *There exist scalars $\epsilon > 0$ and $\tau > 0$ such that*

$$\text{dist}(z \mid Z^*) \leq \tau \left\| z - (z - Mz - q)_+ \right\|$$

for all $z \geq 0$ with $\left\| z - (z - Mz - q)_+ \right\| \leq \epsilon$.

We now give our main theorem. The proof of this theorem is modeled after the proof of Theorem 3.1 given in [11].

Theorem 3 *Suppose that M is symmetric and positive semidefinite and that f given by (21) is bounded from below on $\mathbb{R}^n_+$. Let $\{z^i\}$ be the iterates generated by the matrix splitting algorithm (18), (19), (20). Then $\{z^i\}$ converges at least linearly to an element of $Z^* \bigcap L$.*

Proof We first show that

$$f(z^{i+1}) - f(z^i) \le -\nu/2 \left\| \bar{z}^{i+1} - z^i \right\|^2, \forall i \tag{22}$$

where ν is the smallest eigenvalue of $B - C$. To see this, fix any i. By definition of $\bar{z}^{i+1}$, we have that

$$\left\langle B\bar{z}^{i+1} + Cz^i + q, \bar{z}^{i+1} - z \right\rangle \le 0, \forall z \ge 0 \tag{23}$$

The definition of f leads to

$$f(z^{i+1}) - f(z^i) = \left\langle Mz^i + q, z^{i+1} - z^i \right\rangle + \frac{1}{2}\left\langle z^{i+1} - z^i, M(z^{i+1} - z^i) \right\rangle$$

However, $z^{i+1} = P(\bar{z}^{i+1} \mid L)$ means that

$$\left\langle \bar{z}^{i+1} - z^{i+1}, z - z^{i+1} \right\rangle \le 0, \forall z \in L \tag{24}$$

Since $z^{i+1} - Mz^i - q \in L$ it follows that

$$\left\langle z^{i+1} - \bar{z}^{i+1}, Mz^i + q \right\rangle \le 0$$

Also, $M = B + C$, so that

$$\begin{aligned} f(z^{i+1}) - f(z^i) &\le \left\langle Mz^i + q, \bar{z}^{i+1} - z^i \right\rangle + \frac{1}{2}\left\langle z^{i+1} - z^i, M(z^{i+1} - z^i) \right\rangle \\ &= \left\langle B\bar{z}^{i+1} + Cz^i + q, \bar{z}^{i+1} - z^i \right\rangle + \frac{1}{2}\left\| z^{i+1} - z^i \right\|_M^2 - \left\| \bar{z}^{i+1} - z^i \right\|_B^2 \end{aligned}$$

Noting that $z^{i+1} \pm M(z^{i+1} - z^i) \in L$ it follows from (24) that

$$\left\langle \bar{z}^{i+1} - z^{i+1}, Mz^{i+1} \right\rangle = \left\langle \bar{z}^{i+1} - z^{i+1}, Mz^i \right\rangle$$

Hence

$$\begin{aligned} \frac{1}{2}\left\| z^{i+1} - z^i \right\|_M^2 &= \frac{1}{2}\left\| \bar{z}^{i+1} - z^i \right\|_M^2 + \left\langle z^{i+1} - \bar{z}^{i+1}, M(\bar{z}^{i+1} - z^i) \right\rangle + \frac{1}{2}\left\| z^{i+1} - \bar{z}^{i+1} \right\|_M^2 \\ &= \frac{1}{2}\left\| \bar{z}^{i+1} - z^i \right\|_M^2 + \left\langle z^{i+1} - \bar{z}^{i+1}, M(\bar{z}^{i+1} - z^{i+1}) \right\rangle + \frac{1}{2}\left\| z^{i+1} - \bar{z}^{i+1} \right\|_M^2 \\ &= \frac{1}{2}\left\| \bar{z}^{i+1} - z^i \right\|_M^2 - \frac{1}{2}\left\| z^{i+1} - \bar{z}^{i+1} \right\|_M^2 \end{aligned}$$

Substituting the expression in the inequality above, we get

$$\begin{aligned} f(z^{i+1}) - f(z^i) &\le \left\langle B\bar{z}^{i+1} + Cz^i + q, \bar{z}^{i+1} - z^i \right\rangle + \frac{1}{2}\left\| z^{i+1} - z^i \right\|_M^2 - \left\| \bar{z}^{i+1} - z^i \right\|_B^2 \\ &= \left\langle B\bar{z}^{i+1} + Cz^i + q, \bar{z}^{i+1} - z^i \right\rangle - \frac{1}{2}\left\| z^{i+1} - \bar{z}^{i+1} \right\|_M^2 - \left\| \bar{z}^{i+1} - z^i \right\|_{B-M/2}^2 \\ &\le -\left\| \bar{z}^{i+1} - z^i \right\|_{B-M/2}^2 \\ &\le -\frac{1}{2}\left\| \bar{z}^{i+1} - z^i \right\|_{B-C}^2 \end{aligned}$$

where the second inequality follows from (23) with $z = z^i$ and the fact that M is positive semidefinite and the third inequality follows from $M = B + C$. Thus (22) holds.

Now we claim that

$$\left\| z^i - \left(z^i - Mz^i - q\right)_+ \right\| \leq \kappa_1 \left\| \bar{z}^{i+1} - z^i \right\|, \forall i \tag{25}$$

First, fix an iteration i. From the definition of $\bar{z}^{i+1}$ we have

$$\begin{aligned}
\left\| z^i - \left(z^i - Mz^i - q\right)_+ \right\| &= \left\| z^i - \left(z^i - Mz^i - q\right)_+ - \bar{z}^{i+1} + \left(\bar{z}^{i+1} - B\bar{z}^{i+1} - Cz^i - q\right)_+ \right\| \\
&\leq \left\| z^i - \bar{z}^{i+1} \right\| + \left\| z^i - Mz^i - q - \bar{z}^{i+1} + B\bar{z}^{i+1} + Cz^i + q \right\| \\
&\leq 2\left\| z^i - \bar{z}^{i+1} \right\| + \left\| B(\bar{z}^{i+1} - z^i) \right\| \\
&\leq (2 + \|B\|) \left\| z^i - \bar{z}^{i+1} \right\|
\end{aligned}$$

The first inequality in the above follows from the nonexpansiveness property of projections and the second follows from $M = B + C$. Thus (25) holds with $\kappa_1 = 2 + \|B\|$.

Since, by assumption, f is bounded below on $\mathbb{R}^n_+$, (22) implies that

$$\left\| \bar{z}^{i+1} - z^i \right\| \to 0$$

It then follows from (25) that $\left\| z^i - (z^i - Mz^i - q)_+ \right\| \to 0$ and so by Proposition 2 there exists a scalar constant $\kappa_2 > 0$ and an index $\hat{i}$ such that

$$\operatorname{dist}(z^i \mid Z^*) \leq \kappa_2 \left\| \bar{z}^{i+1} - z^i \right\|, \forall i \geq \hat{i}$$

Thus

$$\operatorname{dist}(z^i \mid Z^*) \to 0$$

It is also well known that f is constant on Z^* so we shall denote this constant value by f^∞.

We now show that $f(z^i) \to f^\infty$ and estimate the speed of convergence. Fix any $i \geq \hat{i}$ and let y^i be defined as follows

$$y^i = (\hat{u}^i_1, \hat{u}^i_1, \hat{u}^i_1)$$

where $\hat{u}^i_1 = P(u^i_1 \mid U^*)$. Then $y^i \in Z^*$ and

$$\left\| y^i - z^i \right\|^2 = 3 \left\| \hat{u}^i_1 - u^1_1 \right\|^2 \leq 3 \operatorname{dist}(z^i \mid Z^*)^2$$

Thus

$$\left\| y^i - z^i \right\| \leq \sqrt{3} \operatorname{dist}(z^i \mid Z^*) \leq \sqrt{3}\kappa_2 \left\| \bar{z}^{i+1} - z^i \right\| := \kappa_3 \left\| \bar{z}^{i+1} - z^i \right\| \tag{26}$$

Now

$$\begin{aligned}
f(z^{i+1}) - f^\infty &= f(z^{i+1}) - f(y^i) \\
&= \left\langle My^i + q, z^{i+1} - y^i \right\rangle + \frac{1}{2} \left\| z^{i+1} - y^i \right\|^2_M
\end{aligned}$$

and $z^{i+1} - My^i - q \in L$, so (24) implies

$$\left\langle z^{i+1} - \bar{z}^{i+1}, My^i + q \right\rangle \leq 0$$

Therefore

$$\begin{aligned}
f(z^{i+1}) - f^\infty &\leq \left\langle My^i + q, \bar{z}^{i+1} - y^i \right\rangle + \frac{1}{2} \left\| z^{i+1} - y^i \right\|_M^2 \\
&\leq \left\langle (B+C)y^i + q, \bar{z}^{i+1} - y^i \right\rangle - \left\langle B\bar{z}^{i+1} + Cz^i + q, \bar{z}^{i+1} - y^i \right\rangle + \frac{1}{2} \left\| z^{i+1} - y^i \right\|_M^2 \\
&= \left\langle C(y^i - z^i), \bar{z}^{i+1} - y^i \right\rangle - \left\| \bar{z}^{i+1} - y^i \right\|_B^2 + \frac{1}{2} \left\| z^{i+1} - y^i \right\|_M^2 \\
&\leq \|C\| \left\| y^i - z^i \right\| \left\| \bar{z}^{i+1} - y^i \right\| \qquad (27)
\end{aligned}$$

the second inequality following from (23) with $z = y^i$ and the last inequality following from the fact that

$$\frac{1}{2} \left\| z^{i+1} - y^i \right\|_M^2 - \left\| \bar{z}^{i+1} - y^i \right\|_B^2 = -\frac{1}{2} \left\| z^{i+1} - \bar{z}^{i+1} \right\|_M^2 - \frac{1}{2} \left\| \bar{z}^{i+1} - y^i \right\|_{B-C}^2$$

and both M and $B - C$ are positive semidefinite. However, by (26)

$$\begin{aligned}
\left\| \bar{z}^{i+1} - y^i \right\| &\leq \left\| \bar{z}^{i+1} - z^i \right\| + \left\| z^i - y^i \right\| \\
&\leq (\kappa_3 + 1) \left\| \bar{z}^{i+1} - z^i \right\|
\end{aligned}$$

which when substituted in (27) and again using (26) gives

$$f(z^{i+1}) - f^\infty \leq \|C\| \, \kappa_3(\kappa_3 + 1) \left\| \bar{z}^{i+1} - z^i \right\|^2$$

Let us define $\kappa_4 := \|C\| \, \kappa_3(\kappa_3 + 1)$. It now follows that

$$\begin{aligned}
f(z^{i+1}) - f^\infty &\leq \kappa_4 \left\| \bar{z}^{i+1} - z^i \right\|^2 \\
&\leq (2\kappa_4/\nu)(f(z^i) - f(z^{i+1}), \forall i \geq \hat{i}
\end{aligned}$$

the last inequality following from (22). If we rearrange terms we find

$$(1 + \frac{2\kappa_4}{\nu})(f(z^{i+1}) - f^\infty) \leq \frac{2\kappa_4}{\nu}(f(z^i - f^\infty), \forall i \geq \hat{i}$$

Hence $f(z^i)$ converges at least linearly to f^∞. By (22), $\{z^i\}$ also converges at least linearly to some z^*. Since $\text{dist}(z^i \mid Z^*) \to 0$, z^* is an element of Z^*. Furthermore, since L is closed if follows that $z^* \in L$. □

We note that Theorem 3 can be extended to cover the case of inexact solution of (18), namely

$$\bar{z}^{i+1} = \left(\bar{z}^{i+1} - (B\bar{z}^{i+1} + Cz^i + q + h^i)\right)_+$$

provided that the error satisfies

$$\left\|h^i\right\| \leq (\frac{\nu}{2} - \epsilon) \left\|\bar{z}^{i+1} - z^i\right\|$$

for some $\epsilon > 0$. The modification of the proof only requires changes to (22) and (27).

We now relate Theorem 3 to our PCD algorithm.

Corollary 4 (PCD Convergence for Quadratic Programs) *Assume that A_l has linearly independent rows and the feasible region of (1) is nonempty. The PCD algorithm defined by (2), (3), (4) and (5) converges linearly, that is*

$$x_l^i \to x^*, l = 1, 2, 3$$

and

$$t_{jl}^i \to s_j^*, j, l = 1, 2, 3, j \neq l$$

with x^* a solution of (1).

Proof As was shown above, the algorithm defined by (2), (3), (4) and (5) is precisely of the form given in (18) and (20) with B defined in (13) and C defined by (14). Hence, in the dual space, the algorithm corresponds to an iterative matrix splitting algorithm followed by a subspace projection. We now show that the conditions required for convergence of this method as given in Theorem 3 are satisfied.

Note that $M = B+C$ and therefore to show that M is positive semidefinite it is sufficient to show (by (16)) that H as defined in (10) is positive semidefinite. This is clear since

$$H = \gamma \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} Q^{-1} \begin{bmatrix} A_1^T & A_2^T & A_3^T \end{bmatrix}$$

and Q is positive definite (hence also Q^{-1}).

We now show $B - C$ is positive definite. Note that

$$B - C = \begin{bmatrix} H + 2J_1 & 0 & 0 \\ 0 & H + 2J_2 & 0 \\ 0 & 0 & H + 2J_3 \end{bmatrix}$$

so we only show that $H + 2J_1$ is positive definite. Suppose that

$$(x_1, x_2, x_3)^T (H + 2J_1)(x_1, x_2, x_3) = 0$$

Then

$$\begin{bmatrix} A_1^T & A_2^T & A_3^T \end{bmatrix} (x_1, x_2, x_3) = 0$$

and $x_2 = x_3 = 0$. Since A_1 has linearly independent rows it now follows that $x_1 = 0$, and so $H + 2J_1$ is positive definite as required. Thus $B - C$ is positive definite and so (19) holds.

It remains to show that f bounded below on positive orthant. The fact that f is bounded below is equivalent to a solution existing by [3]. Since M symmetric and positive semidefinite any solution of

$$\underset{z \geq 0}{\text{minimize}} \quad f(z)$$

solves LCP(M,q) and conversely. As shown above, any solution of LCP(M,q) leads to a solution of LCP(H,h) which is the dual of (1). The fact that this problem has a solution is equivalent to (1) having a solution which by strong convexity is equivalent to (1) being feasible, as assumed above.

Thus by invoking Theorem 3 we see that the dual iterates z^i converge linearly to z^*, a solution of LCP(M,q). Furthermore, $z^* \in L$, so that $z^* = (u^*, u^*, u^*)$. Thus from (9) it follows that

$$t_{jl}^* = s_j^*, \forall l \neq j$$

and so from (7) it follows that

$$\begin{aligned} \bar{x}_l^* &= -Q^{-1}(c + \textstyle\sum_{k=1}^3 A_k^T \bar{s}_k^*) \\ \bar{s}_l^* &= (\bar{s}_l^* + \gamma(A_l \bar{x}_l^* - a_l))_+ \end{aligned} \qquad l = 1, 2, 3$$

Hence $x_l^* = x^*$ for $l = 1, 2, 3$ and we have

$$\begin{aligned} \bar{x}^* &= -Q^{-1}(c + \textstyle\sum_{k=1}^3 A_k^T \bar{s}_k^*) \\ \bar{s}_l^* &= (\bar{s}_l^* + \gamma(A_l \bar{x}_l^* - a_l))_+ \end{aligned} \qquad l = 1, 2, 3$$

But this implies that $\bar{x}^*$ solves (1). □

3 Computational Results

We have tested out the algorithm described above on some linear programming problems. The standard form linear program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \begin{array}{l} Ax = b \\ x \geq 0 \end{array} \end{array}$$

has the dual problem

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y \leq c \end{array} \tag{28}$$

and these problems are in precisely the form of our preceding discussion except the objective is not strongly convex. In order to strongly convexify the objective we have used the least two–norm formulation [7, 6], where for $\epsilon \in (0, \bar{\epsilon}]$ for some $\bar{\epsilon} > 0$, the solution of

$$\begin{array}{ll} \text{minimize} & -b^T y + \frac{\epsilon}{2} y^T y \\ \text{subject to} & A^T y \leq c \end{array} \tag{29}$$

is the least two–norm solution of (28). For the purpose of our computation, a value of $\epsilon = 10^{-6}$ was used.

We have split up the problems as follows: firstly the user has specified the number of processors available and the problem has been split into that many blocks. If the number of constraints in each block is not the same we have added combinations of constraints form other blocks to make the number of constraints in each block equal with the aim of balancing the load in each processor.

The PCD algorithm was implemented on the Sequent Symmetry S–81 shared memory multiprocessor. The subproblems were solved on each processor using MINOS 5.3 a more recent version of [8]. The explicit constraints in each subproblem remained fixed throughout the computation but the blocks were not chosen to satisfy the linear independence assumption.

We have used the following scheme to update the augmented Lagrangian parameter, γ. Initially it is set at 10 and is increased by a factor of 4 only when the norm of the violation of the constraints increases.

The algorithm was terminated whenever the difference in the primal objective value of (28) and its dual objective value normalized by their sum differed by less than 10^{-5}. The constraint violation was also required to be less than this tolerance.

We give two tables below for comparison. Table 1 gives the best results that were obtained using the algorithm described in [2] on the Sequent Symmetry S–81 for 5 small linear programs reformulated as in (29). The first three are homemade test problems, while the last two, AFIRO and ADLittle, are from the NETLIB collection [4]. In the tables, an empty column entry signifies that we did not perform the computation. Note also that these results include a heuristic for calculating a step length.

Table 2 gives the results for the algorithm outlined in this paper. We remark that this algorithm performs uniformly better than the one described in [2]. Furthermore, its implementation is somewhat simpler. Note the strong indication that these results give to the fact that the number of iterations is independent of the number of processors used.

Problem	Variables	Constraints	Blocks			
			3	6	9	18
Ex6	3	5	2	2		
Ex9	5	11	4	5		
Ex10	6	14	4	4	4	
AFIRO	27	51	13	15	15	14
ADLittle	56	138	12	14	14	15

Table 1: Old PCD Algorithm with variable λ

Problem	Variables	Constraints	Blocks			
			3	6	9	18
Ex6	3	5	2	2		
Ex9	5	11	3	3		
Ex10	6	14	3	2	3	
AFIRO	27	51	8	8	8	8
ADLittle	56	138	9	9	9	9

Table 2: New PCD Algorithm with projection step

4 Conclusions

We have presented a method for solving strongly convex quadratic programs with large numbers of linear inequality constraints and have shown the method to be linearly convergent. The method is easy to implement and preliminary computational results are encouraging.

Further extensions of this work are possible when the subspace L is modified appropriately. These extensions will be addressed in a future paper.

References

[1] D.P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods.* Academic Press, New York, 1982.

[2] M.C. Ferris and O.L. Mangasarian. Parallel constraint distribution. Technical Report 971, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin 53706, October 1990.

[3] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 3:95–110, 1956.

[4] D.M. Gay. Electronic mail distribution of linear programming test problems. *COAL Newsletter*, 13:10–12, 1985.

[5] O.L. Mangasarian. *Nonlinear Programming.* McGraw–Hill, New York, 1969.

[6] O.L. Mangasarian. Normal solutions of linear programs. *Mathematical Programming Study*, 22:206–216, 1984.

[7] O.L. Mangasarian and R.R. Meyer. Nonlinear perturbation of linear programs. *SIAM Journal on Control and Optimization*, 17(6):745–752, November 1979.

[8] B.A. Murtagh and M.A. Saunders. MINOS 5.0 user's guide. Technical Report SOL 83.20, Stanford University, December 1983.

[9] R.T. Rockafellar. Augmented Lagrange multiplier functions and duality in nonconvex programming. *SIAM Journal on Control*, 12:268–285, 1974.

[10] R.T. Rockafellar. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operations Research*, 1(2):97–116, 1976.

[11] P. Tseng and Z.-Q. Luo. Error bound and convergence analysis of matrix splitting algorithms for the affine variational inequality problem. Technical Report LIDS-P-1988, MIT, Cambridge, Massachusetts, 1990.