A Relaxed Synchronization Approach for Solving Parallel Quadratic Programming Problems with Guaranteed Convergence

Jyotikrishna Dass, Rabi Mahapatra

Department of Computer Science and Engineering
Texas A&M University, College Station, TX 77843
{dass.jyotikrishna@tamu.edu}

COMPUTER SCIENCE & ENGINEERING
TEXAS A&M UNIVERSITY

Introduction

Main challenge in maximizing processor utilization for distributed computing is to reduce idling due to synchronization across processors. Synchronization is necessary after every iteration , however, it prevents many numerical algorithms from scaling with number of processors.

We relax this requirement by synchronizing at a lower rate, process referred to as Lazy Synchronization. We present a novel approach to solve parallel Quadratic Programming problems using the proposed *Lazily Synchronized Dual Ascent* (LSDA) algorithm. We also provide optimal rate for synchronization which ensures faster convergence to the solution compared to *Tightly Synchronized Dual Ascent* (TSDA) algorithm.

Problem Formulation

$$\min_{x} \sum_{i=1}^{N} f_i(x_i)$$
$$\text{subject to } \sum_{i=1}^{N} A_i x_i = b$$

where, $x_i \in \mathbb{R}^{n/N}$, $A_i \in \mathbb{R}^{m \times \frac{n}{N}}$, $b \in \mathbb{R}^m$, N: number of cluster nodes,

QP Problem:

$$f_i(x_i) = \frac{1}{2} x_i^T Q_i x_i + c_i^T x_i$$

where, $c_i \in \mathbb{R}^{n/N}$ and $Q_i \in \mathbb{R}^{\frac{n}{N} \times \frac{n}{N}}$ is a SPD matrix

LSDA Method:

Step 1: For each cluster node $i=1 \ldots N$,

$$x_i^{k+1} = \underset{x_i}{\operatorname{argmin}}\, L(x_i, y^k) = -Q_i^{-1}(A_i^T y^k + c_i)$$

Step 2: At master node,

$$y^{k+1} = y^k + \eta \sum_{i=1}^{N} (A_i x_i^{tP+1} - \frac{b}{N})$$

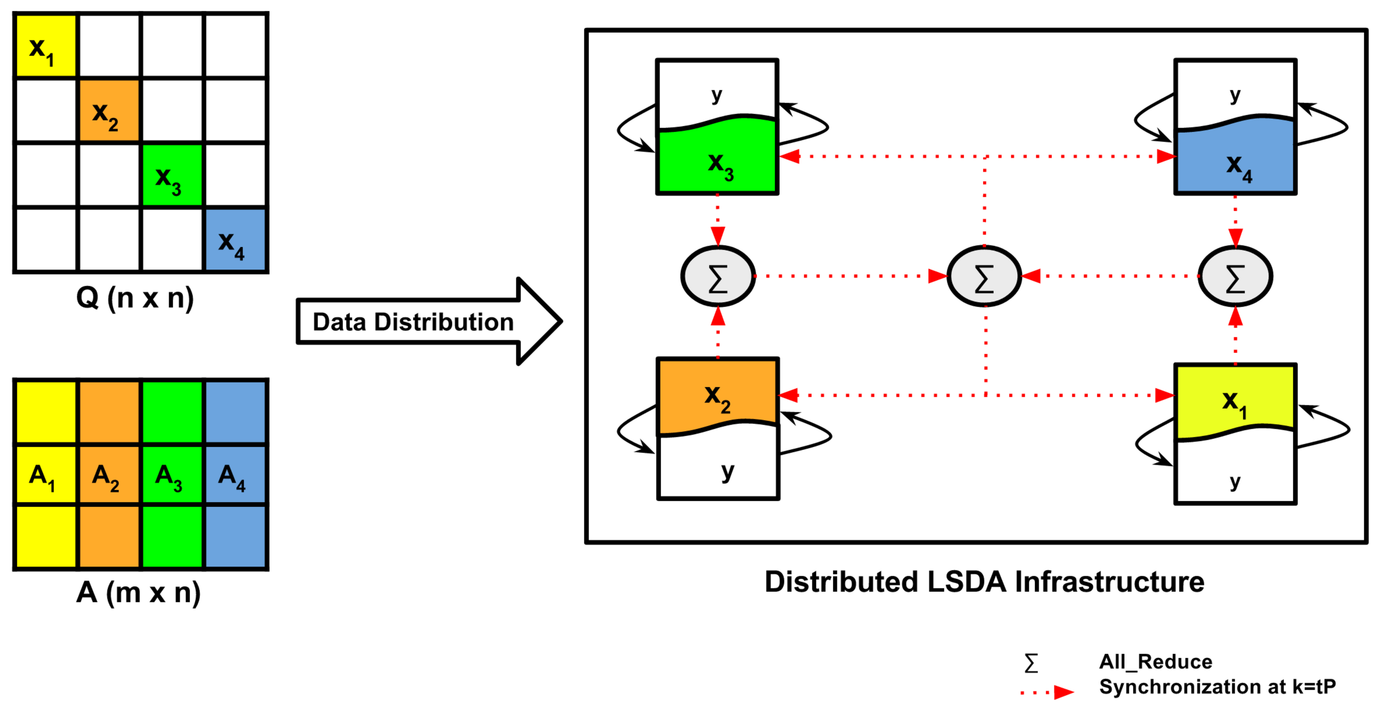
where, L( ) is Langrangian, η is step size, y is dual variable, k is iteration such that, $tP \le k < (t+1)P$, t ∈ and $P \ge 1$ is synchronization period.

Optimal Synchronization Period (P^*):

$$P^* = \max \arg\min_{P \in \mathbb{N}} \max\{|1 - \lambda_{\min}(M)P|, |1 - \lambda_{\max}(M)P|\}$$

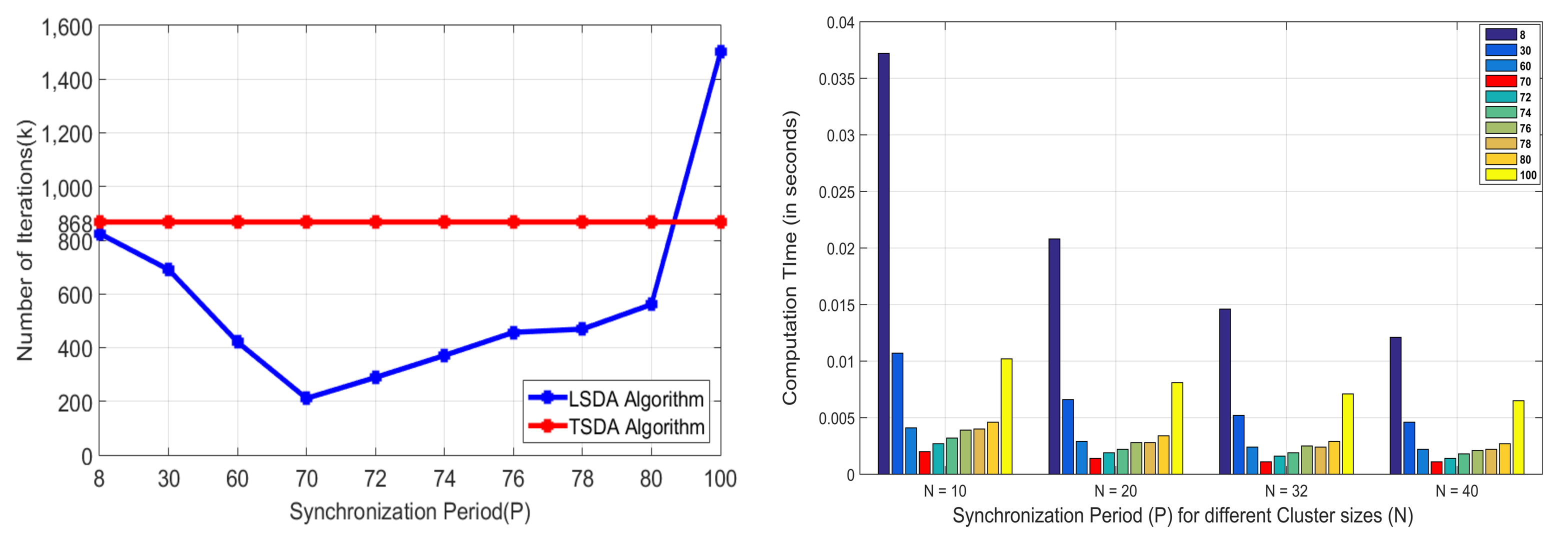
where, $M = \eta \sum_{i=1}^{N} A_i Q_i^{-1} A_i^T$, $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$ denote the minimum and the maximum eigenvalues of the square matrix M, respectively.

LSDA Framework



Results

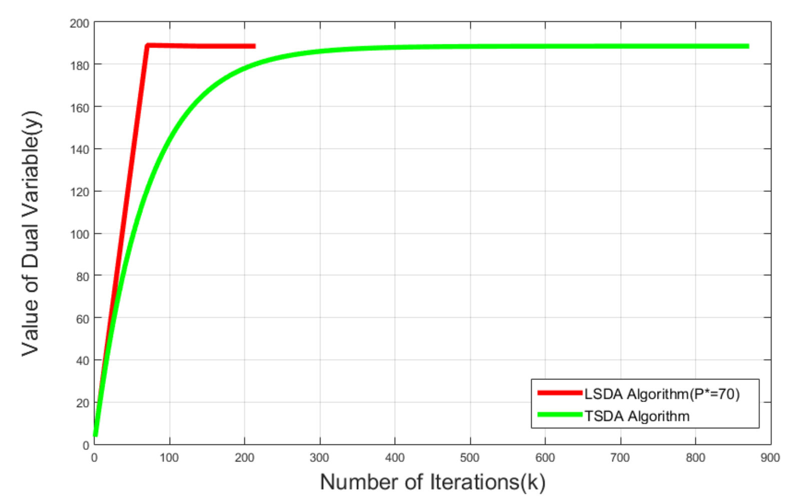
1. Optimal Synchronization Period, *P**



For a given synthetic dataset, minimum number of iterations k required for LSDA to converge occurs at P^* =70, which is independent of the cluster size N. It also validates the analytical formulation for optimal synchronization period.

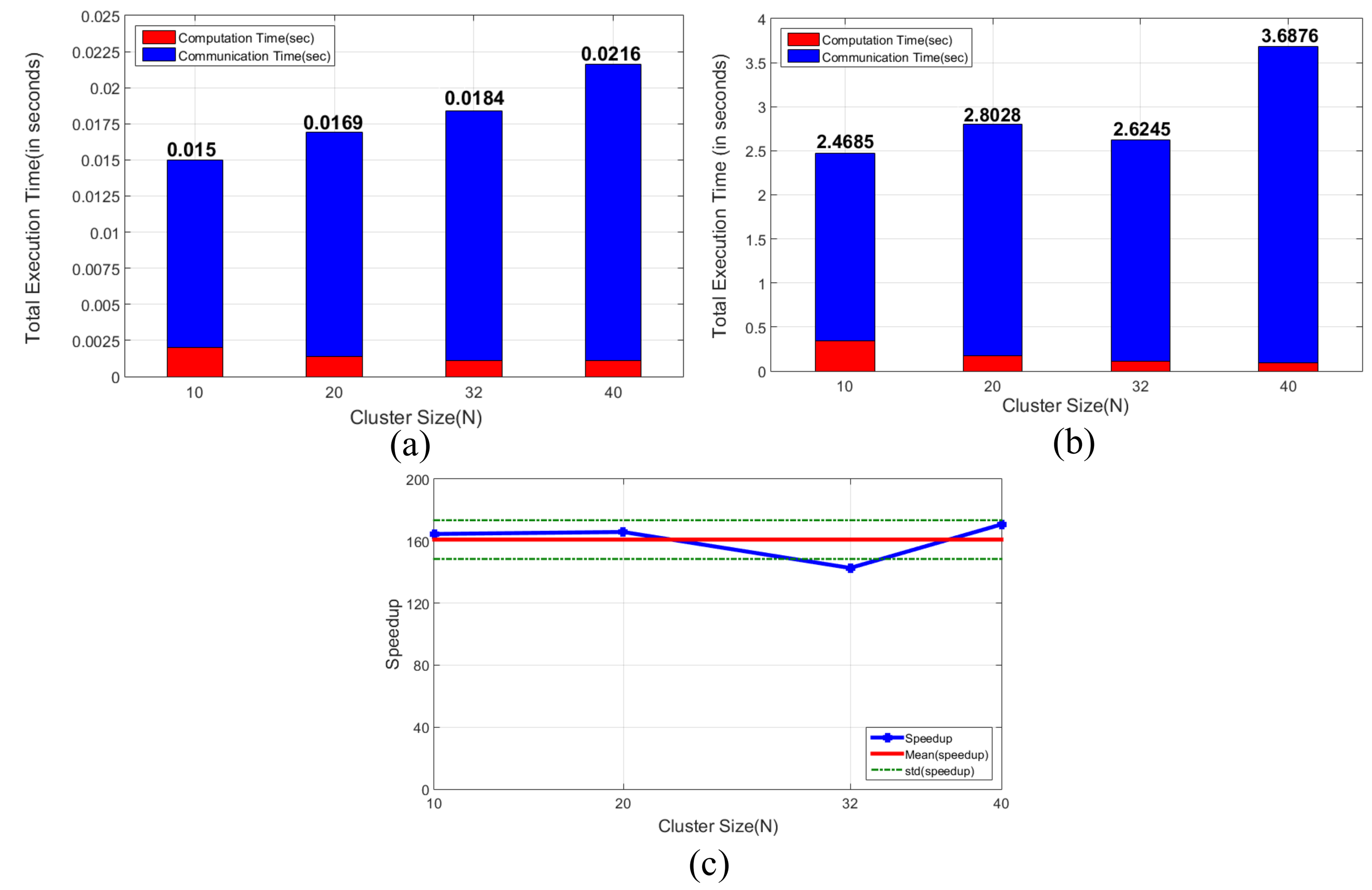
2. Convergence of LSDA vs TSDA

It is observed that LSDA (P^* =70) converges to the optimal solution in k=211 iterations, much faster compared to k=868 iterations for TSDA. In LSDA, synchronization among nodes occurs only thrice (k/P^*) whereas in TSDA it occurs every iteration. Hence, the communication delay is reduced by 99.65% in LSDA.



3. Overall Execution time and Speedup

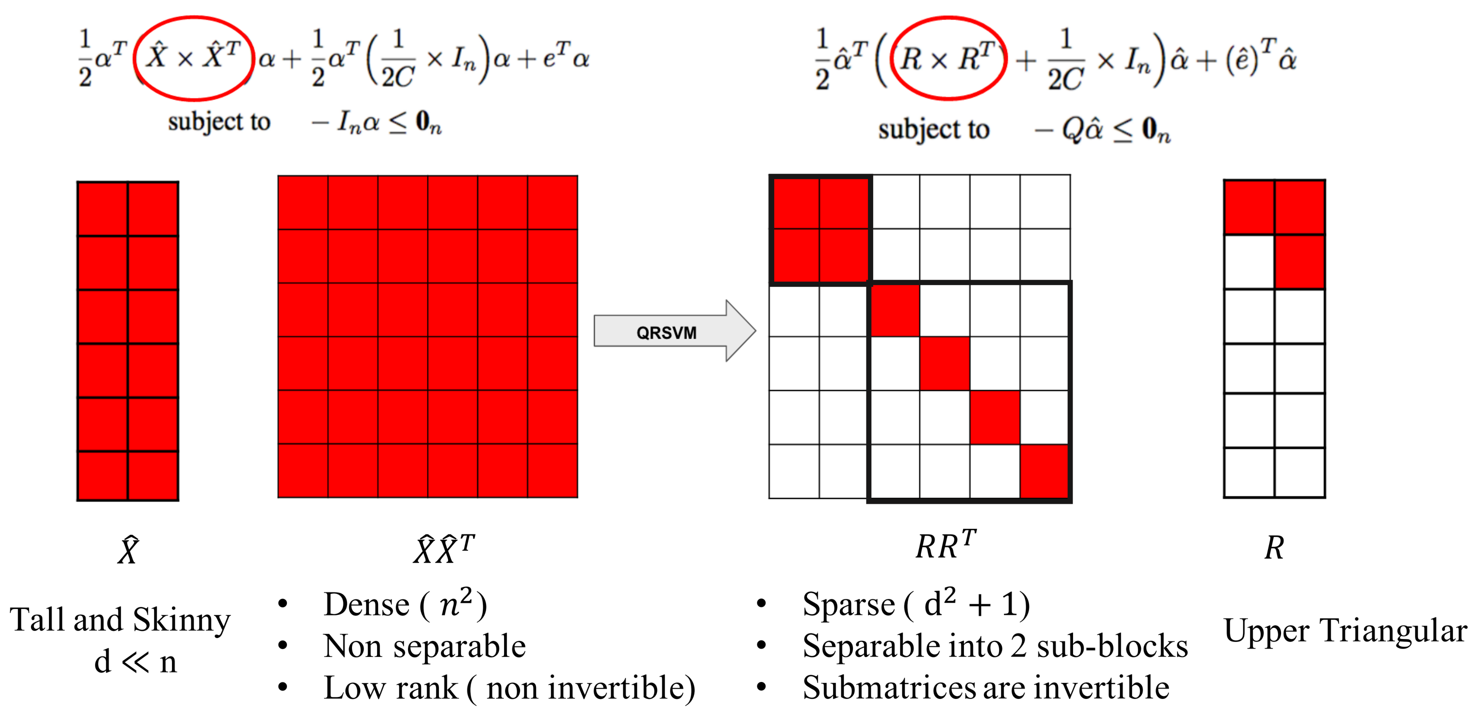
Computation time and communication time (synchronization) , together is defined as the overall execution time. In Figures below, (a) LSDA technique provides significant benefits in both the computation and communication time compared to (b) TSDA method. This is because computationally intensive local x-update calculation and communication demanding synchronization occurs only at optimal synchronization period P^*=70 for LSDA. TSDA performs slower since its synchronization period P^*=1. (c) Speedup of LSDA vs TSDA.

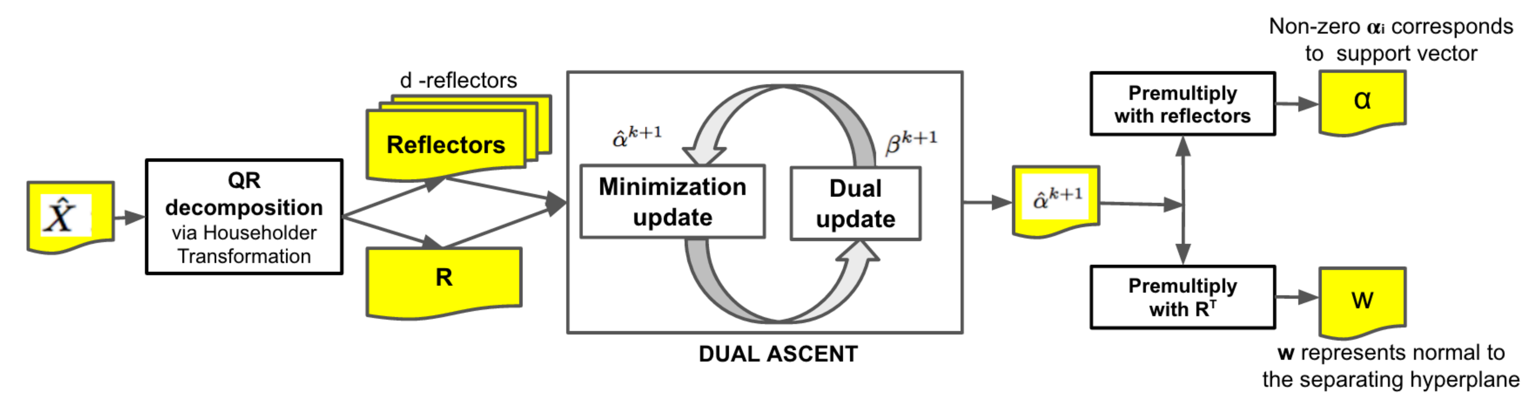


(a) (b) (c)

Current Research

QRSVM: QR decomposition framework to solve large-scale linear *Support Vector Classification* problems having huge number of instances with relatively smaller dimensionality, i.e. $d \ll n$.





QRSVM framework comprises of two main stages, namely, 1) *QR decomposition* of the original input matrix X̂ into Householder reflectors and a matrix R, and 2) *Dual Ascent method* to solve the QRSVM problem for obtaining the normal w to the hyperplane and identifying set of support vectors.

Acknowledgements

1. **LSDA:** Dr. Kooktae Lee, Dr. Raktim Bhattacharya and Prithvi Sakuru
2. **QRSVM:** Prithvi Sakuru and Dr. Vivek Sarin
3. TAMU Supercomputing Facility